/510R67M>C1/

23950
Root

St. Procopius
College Library
LISLE, ILL.

Borrowers are responsible for all books taken out on their cards.

Books may be kept for two weeks and renewed once for the same period, except fiction, 7-day, and reserved books.

A fine of one cent per day will be charged on each book which is not returned according to the above rule.

Digitized by the Internet Archive
in 2025

https://archive.org/details/bwb_T5-BYJ-864

THE MATHEMATICS OF ENGINEERING

THE MATHEMATICS OF ENGINEERING

By

RALPH E. ROOT, PH.D.

Professor of Mathematics and Mechanics,
Post Graduate School,
United States Naval Academy

WITHDRAWN

BALTIMORE
THE WILLIAMS & WILKINS COMPANY
1927

510
R67m

COPYRIGHT 1927
THE WILLIAMS & WILKINS COMPANY

Made in United States of America

Published June, 1927

COMPOSED AND PRINTED AT THE
WAVERLY PRESS
FOR
THE WILLIAMS & WILKINS COMPANY
BALTIMORE, MD., U. S. A.

23950

PREFACE

This book is an outgrowth of the efforts of the author, during about thirteen years, to meet the mathematical requirements of student officers at the Postgraduate School of the Naval Academy. First the material of certain chapters appeared to supplement a standard work on calculus, then other chapters appeared to replace the standard text, the work growing chapter by chapter as the need arose. Since 1923 mimeographed copies have been in use at the Postgraduate School, essentially in the present form. Various groups of our student officers are to specialize in different branches of engineering, as mechanical, electrical, civil and aeronautical, and in radio communication, aerology, naval construction and several branches of ordinance engineering. The final stage of instruction for each group consists of one or more years of advanced technical instruction at a university or technical school preeminent in the specialty of the group concerned. It is believed, therefore, that if we have properly estimated the requirements in the local situation we have, of necessity, taken cognizance of some of the best tendencies in engineering education and have approximated fairly well the mathematical requirements of the modern engineer.

While it may appear to some that such material as that of Chapters V and VII should be unnecessary for students of sufficient maturity to pursue other portions of the book, the author has found these chapters helpful in covering, in a minimum of time, certain essential reviews. In the hands of an able teacher Chapter VII might even serve for a first course in analytic geometry, but, like Chapter V, it is offered as a brief, condensed restatement of principles and methods which the student may have covered previously. Some may wish to reserve these chapters for reference only. The idea of the condensed restatement extends, in some degree, to the earlier chapters on differentiation and integration, as most students to whom this book may appeal will probably have had, like our own students, an introductory course in calculus.

Concerning the more advanced topics treated, we must admit that this work stops far short of the complete treatment that a mathematician might find more to his taste. We have been more anxious to include topics of a sufficiently wide range, and to give to each the

treatment best suited to the needs of the engineer, always bearing in mind the brevity essential to a single course of reasonable proportions.

In the matter of rigor in the statement and proof of propositions we cannot lay claim to consistency. We have undertaken to make careful and precise statements where this has seemed essential to adequate understanding and safe use of the principles. In the elementary portions certain proofs have been given in detail, others only sketched in outline, some hinted at and others omitted entirely. In the advanced portions more attention has been given to explanations which promote correct understanding than to the development of mathematical rigor. Here also, when brevity demanded it and intelligent use could thus be conserved, certain results have been stated without proof.

In such a work as the present one much must necessarily be left to the teacher. The treatment is based on the supposition that the student is to benefit by class room discussions. The independent reader not thoroughly familiar with the conventional courses in mathematics, at least the freshman courses, will find it helpful to have at hand standard text books on algebra, trigonometry and analytic geometry. For one who would pursue certain of the more advanced topics further we may suggest such works as Wilson's "Advanced Calculus," Osgood's "Advanced Calculus," the Goursat-Hedrick "Mathematical Analysis," works on probability and least squares, on Fourier series and harmonic analysis, and on differential equations. Book lists have been appended at the ends of certain chapters to aid the student in selecting material for further reading, either to amplify the topics of the chapter or to extend the study to other allied topics.

While this book is written for the engineer, and is planned to be convenient and useful for reference, it is not in any sense a handbook. It is supposed that the student will have at hand some engineering handbook, and it is absolutely essential that he should have access to certain mathematical tables. In particular, numerical tables of logarithms, of trigonometric and exponential functions and of the probability integral, as well as a brief table of integrals, are certainly indispensable. Other tables, as hyperbolic functions, natural logarithms, elliptic integrals, etc., are extremely desirable.

To maintain the engineer's point of view and to hold his interest, certain illustrative material has been drawn from technical subjects, but there has been no desire to encroach upon any engineering field further than to suggest the practical bearing of the mathematical topic in hand. Many of the exercises are chosen to suggest such practical bearing, but they are so stated, in general, that the student may in-

telligently apply the mathematical principles without going aside to seek technical knowledge. Our own students are taking courses in mechanics and engineering subjects at the same time that they take the course in mathematics. It is regarded as highly desirable that mechanics be undertaken at an early period in the mathematics course, and that some fundamental engineering courses parallel its later progress.

We believe that the time required to cover this book with reasonable thoroughness should normally be about 125 or 150 lessons. This would vary with students of various degrees of maturity, and may be shortened by omission of certain elementary material. Of course no student will perform all of the exercises, but it is earnestly recommended that each student at least read all exercises, since they teach much that would otherwise be lost, and since, in some cases, they are progressively graded, leading from the simpler ideas to those more involved.

While many of the exercises are original with the author, others have been drawn from many different sources during the years that the book has been in preparation. It is impossible at this time to enumerate those sources or to give credit where it may be due. The author is indebted to his colleagues for helpful criticisms and suggestions, particularly to Professors C. C. Bramble and C. H. Rawlins, who have used the mimeographed text for several years. To these and to student officers of present and former years he is indebted for aid in elimination of errors in the text and exercises. Finally we acknowledge our obligation to present and former heads and administrative officers of the Postgraduate School for their encouragement and cooperation, and for their interest in transmitting constructve criticisms of former student officers who have completed their special courses.

Ralph E. Root.

Annapolis, Md., 1927.

CONTENTS

CHAPTER I

Functions, Notations, Graphs

Art. 1. Functional relations and notations 1
Art. 2. Classification of functions 3
Art. 3. Graphs of simple functions 5
Exercises I 14
Art. 4. Graphs of compound functions 15
Exercises II 19

CHAPTER II

Limits and Continuity

Art. 5. Limits of functions 21
Exercises III 26
Art. 6. Unbounded variation 28
Exercises IV 32
Art. 7. Continuity of functions 34
Exercises V 37

CHAPTER III

Derivatives of Functions

Art. 8. Meaning of derivatives 39
Exercises VI 44
Art. 9. General rules for finding derivatives 45
Art. 10. Derivatives of algebraic functions 49
Exercises VII 51
Art. 11. Derivatives of trigonometric and inverse trigonometric functions 53
Exercises VIII 57
Art. 12. Derivatives of exponential and logarithmic functions 58
Exercises IX 61
Art. 13. Functions defined indirectly 62
Exercises X 65
Art. 14. Successive derivatives 66
Exercises XI 69

CHAPTER IV

Integrals of Functions

Art. 15. Inverse of differentiation 70
Exercises XII 73

Art. 16. Formulas for integration 74
Exercises XIII 77
Art. 17. The definite integral 79
Exercises XIV 81
Art. 18. Limit of a sum of differentials 82
Exercises XV 86

CHAPTER V

Solution of Equations

Art. 19. Algebraic equations, exact roots 88
Exercises XVI 95
Art. 20. Determinants and elimination 96
Exercises XVII 105
Art. 21. Equations that require trigonometric transformation 107
Exercises XVIII 113
Art. 22. Equations involving logarithmic transformation 114
Exercises XIX 115
Art. 23. Approximate solution of equations 115
Exercises XX 123

CHAPTER VI

The Process of Integration

Art. 24. Integration by substitution 125
Exercises XXI 127
Art. 25. Integration by parts 128
Exercises XXII 130
Art. 26. Use of trial integrals 130
Exercises XXIII 132
Art. 27. Integration of rational fractions 132
Exercises XXIV 137
Art. 28. Limitations of the process of integration 137
Exercises XXV 144
Art. 29. Use of tables of integrals 144
Exercises XXVI 146

CHAPTER VII

Certain Standard Curves, Equations, Transformations

Art. 30. Certain uses of coordinates 148
Exercises XXVII 156
Art. 31. The straight line 157
Exercises XXVIII 162
Art. 32. Transformation of coordinates 163
Exercises XXIX 167
Art. 33. The conic sections 168
Exercises XXX 178

Art. 34. The equation of the second degree 179
Exercises XXXI 184
Art. 35. Certain families of curves and lines 184
Exercises XXXII 190

CHAPTER VIII

Certain Uses of Derivatives

Art. 36. Tangents and normals to curves 192
Exercises XXXIII 195
Art. 37. Use of graphs of derived functions 197
Exercises XXXIV 202
Art. 38. Maxima and minima 203
Exercises XXXV 209
Art. 39. Osculating circle, curvature 210
Exercises XXXVI 215

CHAPTER IX

Applications of Integration

Art. 40. Length of arc of a curve 217
Exercises XXXVII 219
Art. 41. Areas and volumes 220
Exercises XXXVIII 226
Art. 42. Moments and mean values 227
Exercises XXXIX 232
Art. 43. Distributed quantities; masses, forces 233
Exercises XL 240
Art. 44. Successive integration 241
Exercises XLI 248
Art. 45. Substitutes for integration 249
Exercises XLII 255

CHAPTER X

Methods of Evaluating Functions

Art. 46. Indeterminate forms 258
Exercises XLIII 266
Art. 47. Taylor's theorem with remainder 267
Exercises XLIV 271
Art. 48. Infinite series 272
Exercises XLV 279
Art. 49. Operations on infinite series 281
Exercises XLVI 288

CHAPTER XI

Complex Quantities and Periodic Functions

Art. 50. Complex quantities 290
Exercises XLVII 302

Art. 51. Simple harmonic functions 303
Exercises XLVIII 310
Art. 52. Analysis of periodic functions 311
Exercises XLIX 324

CHAPTER XII

Functions of Several Variables

Art. 53. Geometrical considerations 327
Exercises L 336
Art. 54. Partial and total derivatives 337
Exercises LI 345
Art. 55. Applications of partial and total derivatives 346
Exercises LII 355
Art. 56. Certain properties of plane curves 357
Exercises LIII 364
Art. 57. Integrals in several variables 365
Exercises LIV 382

CHAPTER XIII

Treatment of Empirical Data

Art. 58. Theory of probability 385
Exercises LV 392
Art. 59. Method of least squares 393
Exercises LVI 406
Art. 60. Practical bearing of precision 408
Exercises LVII 416
Art. 61. Empirical equations 418
Exercises LVIII 424

CHAPTER XIV

Ordinary Differential Equations, First Order

Art. 62. Meaning of differential equations 427
Exercises LIX 430
Art. 63. Separation of variables 431
Exercises LX 435
Art. 64. Exact differential equations 436
Exercises LXI 440
Art. 65. Equations of higher degree 441
Exercises LXII 448

CHAPTER XV

Ordinary Differential Equations, Higher Order

Art. 66. Homogeneous linear equations, constant coefficients 449
Exercises LXIII 452

Art. 67. Non-homogeneous linear equations 454
Exercises LXIV 462
Art. 68. Solution by reduction of order 464
Exercises LXV 467

CHAPTER XVI

Ordinary Differential Equations, Several Variables

Art. 69. Systems of simultaneous equations 469
Exercises LXVI 476
Art. 70. Total differential equations 479
Exercises LXVII 486

CHAPTER XVII

Partial Differential Equations

Art. 71. Meaning of partial differential equations 490
Exercises LXVIII 494
Art. 72. Partial differential equations of first order 495
Exercises LXIX 501
Art. 73. Partial differential equations of higher order 501
Exercises LXX 511

CHAPTER XVIII

Differential Equations, Additional Devices

Art. 74. Approximate solution of differential equations 514
Exercises LXXI 528
Art. 75. Solution by series 529
Exercises LXXII 533

CHAPTER I

FUNCTIONS, NOTATIONS, GRAPHS

Art. 1. Functional Relations and Notations

The student is familiar with the idea of quantity, and understands that a quantity may have different **values,** and that these values may be expressed either as an abstract number or in terms of certain conventional units. A quantity that holds a fixed value throughout a discussion is called a **constant,** while a quantity that takes, or is assigned, different values during a discussion is called a **variable.** To illustrate, the quantity of water in a tank that is being emptied through an orifice may be expressed in cubic feet, cubic inches, pounds, tons, gallons, barrels, etc., according to our convenience. In applying the mathematical operations to such a quantity it is convenient to drop the denomination of the unit, operate with the abstract number, and restore the proper denomination at the end of the process. If we are discussing the time required to empty the tank, or the rate at which the surface of the water lowers, we must regard the water as a variable quantity, while the dimensions of the tank are regarded as constant. If we are discussing the pressure at various points in the tank when the water surface is at a certain height, we must regard the water as a constant quantity, while we consider the pressure at **varying** depths below the surface.

Variable quantities, and constant quantities whose values it is not desired to specify, are denoted by letters or symbols. A variable y may be so related to a variable x that there is a correspondence between the values of y and the values of x. If a single value of y corresponds to each value of x, then y is a **single valued function** of x. If two or more values of y corresponds to the same value of x, y is a **double valued** or **multiple valued function** of x. In either case y is said to be a function of x, i.e., a **functional relation** exists. In the above illustration, if the dimensions of the tank and the size of the orifice are fixed, the time required to empty the tank is a function of the depth of water at the start; if the initial depth of water and the dimensions of the tank are fixed, the time is a function of the diameter of the orifice; if the initial depth of water and the size of orifice are fixed, and the tank is a vertical cylinder, the time is a function of the radius of the cylinder.

Thus it is recognized that the time required to empty the tank is a function of the initial depth of water, the size of the orifice and the dimensions of the tank. Most quantities that require our attention in practice are in reality functions of several variables, but when all but one of these variables are regarded as holding fixed values, we may give attention to the relation between the quantity and the remaining variable, regarding it as a function of one variable.

Functional relations may be defined or specified in various ways. Perhaps the simplest way is to specify a rule for computing the value of y that is to correspond to any given value of x. Following are instances of such rules:

$$y = 2x + 3;\ y = \sqrt{6 - x^2};\ y = \frac{x^2 - 4}{x + 3}.$$

Another method is to specify a geometrical or physical relation between y and x. We may let y be the volume of a sphere and x the radius, or y the deflection of a certain beam and x the load applied, or y the velocity of flow of water from a certain pipe and x the "head" of water. Still another way to define a function is to give corresponding values of x and y in tabular form. Instances of this are the logarithmic and trigonometric tables with which the student is familiar. A fourth method of specifying a functional relation is to present a chart or graph by means of which the value of y, for any given value of x, may be determined. When it is desired to speak of functional relations without specifying the relation, or when the function is specified and it is desired to substitute a convenient symbol for the function, the notation usually employed is:

$$y = f(x), \qquad y = F(x), \qquad y = \phi(x), \text{ etc.}$$

We read, y equals the function f of x, or y equals f of x; y equals F of x; y equals ϕ of x. The different letters, f, F, ϕ, etc., distinguish between different functions.

When we are regarding a variable y as a function of a variable x, the variable x is called the **independent variable,** and y is called the **dependent variable.** The notion of function involves two essential elements; (1) The set of values that the independent variable may take, called the **range** of the independent variable. (2) A law of correspondence, or functional relation, by which a value is assigned to the dependent variable for each value in the range of the independent variable. The variables have notations, but they have no characteristics of their own apart from the range and the law of correspondence.

If y is a function of x, then there is a certain range of values of x for which y takes definite corresponding values. The set of values that y assumes may be regarded as a range for y, and x may be regarded as a function of y on this range, the value of x for any given value of y being that value to which this value of y corresponds when y is regarded as a function of x. Thus the dependent and independent variables may be reversed. It will be sufficiently obvious that y may be a single valued function of x while x is a multiple valued function of y.

Art. 2. Classification of Functions

Functions may be classified according to the nature of the range of the independent variable. Thus we have:

(1) Functions of a **discrete** variable. For example, the sum of a geometrical progression whose first term is a and in which the common ratio is r is a function of the number of terms, n,

$$s = \frac{a(1 - r^n)}{1 - r}$$

The independent variable n takes the **discrete** values, 1, 2, 3, etc., having for a range the set of positive integers. Other instances of such functions are the population of a city as a function of the time in years, and the distance a post or pile is driven by the last blow as a function of the number of blows struck.

(2) Functions of a **continuous** variable. This is by far the more important case. In the progression above, if a and n are fixed, s is a function of the variable r, which may take any real value whatever. No real values being excluded from the range of r, i.e., there being no breaks or gaps in its range, it is said to be a **continuous** variable. In the functions.

$$y = \sin^{-1}x, \qquad y = \sqrt{4 - x^2}, \qquad y = \log x$$

the variable x is a continuous variable, but if we wish to restrict attention to real values of y the range of x in the first case is from -1 to $+1$, inclusive, in the second from -2 to $+2$ inclusive, while in the third case it includes all positive real values.

The term **real value** in the last paragraph is used in opposition to **complex value,** i.e., a value involving an even root of a negative number. This suggests distinction between functions of a **real variable** and functions of a complex variable. The functions of the latter class are the subject of study in an important branch of higher mathe-

matics, and in the present work we give some attention to complex quantities but the functions we shall study are for the most part associated with independent variables whose ranges are confined to real numbers.

A more important classification of functions is concerned with the character of the law of correspondence. Functions are grouped according to the mode of determining the value of the function for a given value of the independent variable.

(1) **Algebraic functions.** In these the value of y may be derived from the value of x by a finite number of arithmetical operations. Algebraic functions are designated as:

(a) **Rational algebraic functions,** which may involve addition, subtraction, multiplication and division. A special class of these is **Polynomials** which are formed by addition, subtraction and multiplication but not division. The process of addition, subtraction, multiplication and division are known as the "rational processes." The following are examples of this class of functions:

$$x^3 - \frac{5}{3}x + 9 \text{ (Polynomial)}, \frac{x-2}{x+5}, \quad x^4 - \frac{6}{x}.$$

(b) **Irrational** functions, which involve the further operation of extracting roots, as

$$\sqrt{x+1}, \sqrt[3]{1-x^2}, x + \sqrt{1-x}, \frac{x^2}{\sqrt{1-x}}, \text{ etc.}$$

(2) **Transcendental functions.** These are functions that are not algebraic, there being no finite number of operations to obtain their values. They are definitely defined in mathematical terms, however, and their definitions yield processes by which their values may be approximated to any degree of accuracy desired. They are usually given names and specific notations, and their values are given in tabular form. Important classes of transcendental functions are:

(a) **Trigonometric** functions, as sin x, cos 2x, tan 1/2x, etc.
(b) **Inverse trigonometric** functions, as $\sin^{-1}x$, $\sec^{-1}x$, etc.
(c) **Exponential** functions, as 10^x, e^{2x}, etc.
(d) **Logarithmic** functions, as log x, log $(1 - x^2)$, etc.

(3) **Empirical functions.** This term is used here to designate a very important class of functions whose definitions cannot be formulated in mathematical terms. The functional relation or law of correspondence

is defined in the language of some science, physical, medical, social, economic, etc. The values of the function for various values of the independent variable are determined by measurement, observation, or count, usually only approximately. Corresponding values of the dependent and independent variables are recorded in tabular or graphical form, and in some cases algebraic or transcendental functions are found which agree so closely with the empirical functions that they are substituted for them for purposes of mathematical manipulation. This class of functions receives special consideration in Chapter XIII.

Art. 3. Graphs of Simple Functions

We now consider the problem of constructing the graph of a function that is given in some form other than by a graph. We may suppose that the values of the function are given in tabular form, or else that a law is stated by means of which values of the function may be determined for values of the independent variable. If the function values are not given in tabular form, it will usually be advantageous to construct such a table of values, corresponding to suitably chosen values of the independent variable, as a step toward making the graph. The next step is to choose a scale for the independent variable usually laid off on a horizontal base line, called the **axis of abscissas,** the value of the unit being so chosen that the range for the variable, or as much of it as it is desired to consider, occupies a suitable length of line on the paper to be used. The third step is to select a scale for the function values usually laid off on a vertical base line, called the axis **of ordinates,** the unit being so chosen, in general, that extreme values of the function are represented on the paper. In most cases there will be no significant relation between values of units in the two variables, since the quantities they represent are of different kinds, but in certain geometrical problems it is desirable to keep the vertical and horizontal units equal. The fourth step is to represent by a point on the paper each pair of corresponding values in the table constructed, the point being set opposite those positions, in the horizontal and vertical scales that belong to the values represented. If the values of the independent variable have been suitably chosen, it should now be possible to complete the graph by joining together in some manner the points so located. The terms **abscissa** and **ordinate** are applied to values measured horizontally and vertically, respectively.

These points will now be illustrated and amplified by consideration of the several classes of functions enumerated in Art. 2. We may dismiss the case of the discrete variable, however, with the statement

that every value that the independent variable may assume is usually indicated on the horizontal scale, and the corresponding points of the graph are either left isolated, or successive ones are joined by straight line segments to complete the graph. Our present interest attaches primarily to functions of the continuous variable, and we proceed to take up such functions under the second classification.

(1) **Algebraic functions.** Consider $y = f(x)$, where $f(x)$ is an algebraic function of x. As we have seen, the range of values for either variable may be restricted to a part of the real numbers, or may extend indefinitely, and interest may attach to negative as well as positive values. Construction of a table of values will reveal the features of special interest, and will usually enable us to choose scales of x and y that interpret these features in the graph.

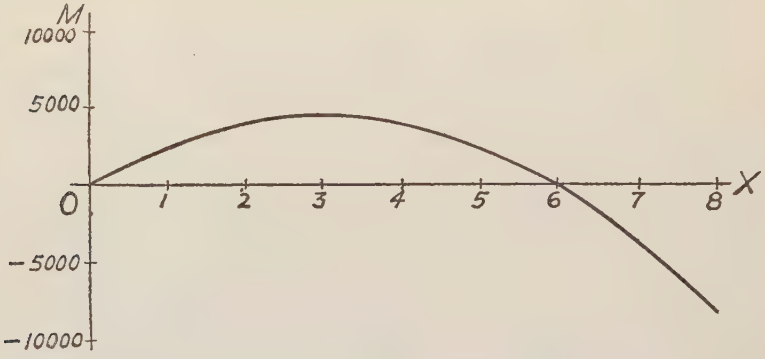


Fig. 1

To illustrate, consider a beam eight feet long, fixed at one end and supported at the other, loaded with 1000 pounds per foot run. The bending moment is given by the equation

$$M = 3{,}000x - 500x^2,$$

where x is distance in feet from the supported end, and M is in pounds-feet. Making a table of values, we observe that within the part of the x range that interests us, i.e., from 0 to 8, the range of M is from −8000 to 4,500. It is a simple matter to choose scales so that the graph reveals the character of the function. The points located are joined by a smooth curve, since it may be presumed that the computed value of M for any intermediate value of x would give a point on such a curve.

x.....	0	1	2	3	4	5	6	7	8
M.....	0	2500	4000	4500	4000	2500	0	−3500	−8000

As another example, consider the velocity v of a ball thrown upward with an initial velocity of 64 feet per second, as a function of its distance above its starting point. Taking the acceleration due to gravity as 32 feet per second per second downward, we have

$$v = \sqrt{v_0^2 - 2gh} = \sqrt{4096 - 64h},$$

where v is velocity in feet per second and h is distance in feet above the starting point. Convenient values for h are multiples of 8, and we find that 64 is the largest value for h, while large negative values for h are without interest. Since v changes rapidly for values of h between

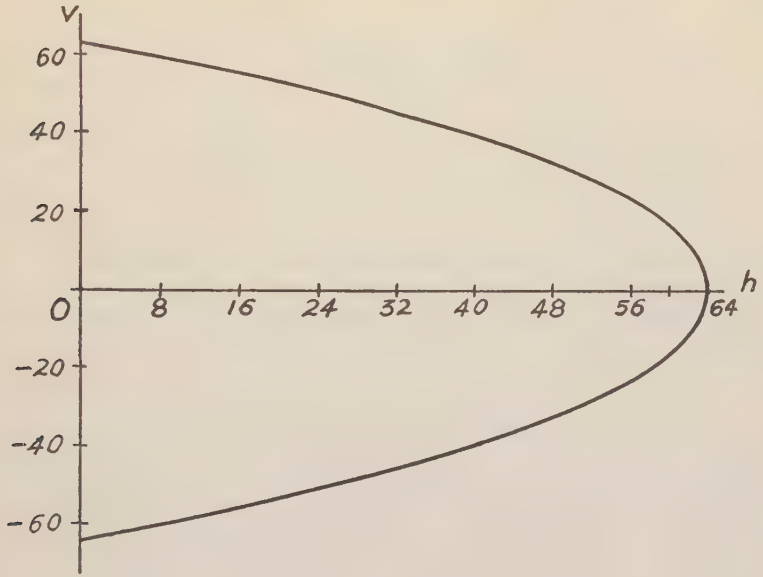


Fig. 2

56 and 64, we compute one intermediate value at h = 60. The values at h = 8, 24, and 40 might have been omitted from the table without material loss of accuracy. The velocity is a double valued function of h, the negative value being velocity of the ball when at the height h coming down. The choice of scales offers no difficulty.

h.....	0	8	16	24	33	40	48
v.....	±64	±59.8	±55.4	±50.6	±45.1	±39.2	±32.0
h.....	56	60	64	−8			
v.....	±22.6	±16.0	0.0	±67.7			

(2) **Transcendental functions.** We confine attention here to the four important classes of transcendental functions previously enumerated.

(a) **Trigonometric functions.** The independent variable here may be regarded as an angle, but in all numerical work in which the magnitude of this quantity is compared to other magnitudes, the unit of the angle is the radian, and it is usual to regard the quantity as an abstract number. In practical instances of the use of such functions as sin x, tan x, etc., the quantity x is as likely to be time, distance, etc., as it is to be an angle. For example, the approximate formula for the angular position of a pendulum is

$$\theta = \theta_0 \cos\sqrt{\frac{g}{L}}\ t.$$

where θ is the angle the pendulum makes with the vertical, θ_0 is the initial value of θ, g is the acceleration due to gravity, L is the length of the pendulum, and t is the time in seconds since the motion started. Here we have the cosine of a quantity involving g, L and t, none of which is an angle. For given values of these three quantities we have the cosine of a certain number, and the value of the cosine is found by interpreting the number as so many radians, which may be converted to degrees by the relation

$$180^\circ = \pi = 3.1416.$$

To illustrate the method of making a graph, let us plot θ as a function of t, taking $\theta_0 = 10^\circ$, and $L = \frac{1}{16}g$. We have

$$\theta = 10^\circ \cos 4t.$$

It will be convenient, in making our table of values, to take t so that 4t is a multiple of $\frac{\pi}{6}$ or 30°. Since the function repeats after 4t reaches 2π, and since the values from π to 2π are negatives of those from 0 to π, we tabulate only the values from 0 to π, corresponding to one swing of the pendulum. For convenience we tabulate corresponding values of t, 4t, cos 4t and θ, which are easily determined in the order named. The scale for t is chosen so that a complete oscillation of the pendulum is shown, and the scale for θ is taken a convenient magnitude to show the character of the curve.

t.....	0	$\pi/24$	$\pi/12$	$\pi/8$	$\pi/6$	$5\pi/24$	$\pi/4$
$4t$.....	0	$\pi/6$	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π
$\cos 4t$.....	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{1}{2}\sqrt{3}$	-1
θ.....	10°	8.66°	5°	0°	−5°	−8.66°	−10°

The values of the trigonometric functions are known at multiples of $\pi/6$, or of $\pi/4$, without use of trigonometric tables. If it is desired

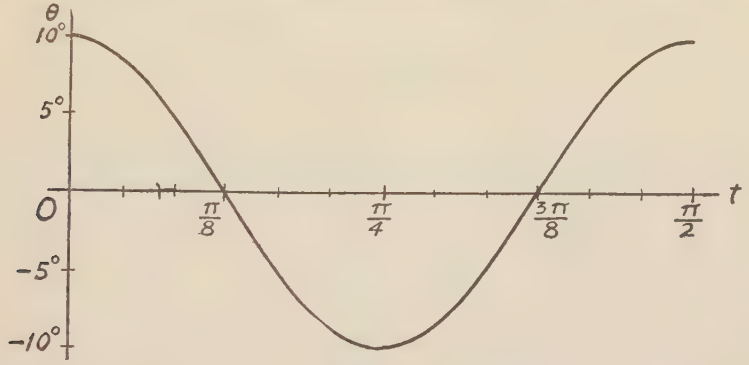


Fig. 3

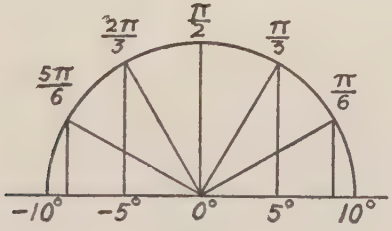


Fig. 4

in any problem to use intermediate values the tables may be used, or the geometrical definitions of the functions may be used to determine the values graphically. For example, in making the graph (Fig. 3) we may draw a circle whose radius is equal to the coefficient of cos 4t, i.e., 10°, in the scale chosen, and, marking the value of 4t on the circumference of the circle the value of θ is the projection of the corresponding radius on the initial line. The values of θ are marked on the initial line. The construction is shown for one swing of the pendulum, i.e., for 4t from 0 to π (Fig. 4).

The graph of any trigonometric function may be constructed by

methods similar to the foregoing. All are periodic, i.e., they repeat their values in cycles as the independent variable increases in value, so that one cycle is sufficient for the graph.

(b) **Inverse trigonometric functions.** It is usually convenient to transform inverse trigonometric functions so as to deal with trigonometric functions instead of plotting directly. If we have

$$y = 5 \tan^{-1} 16x,$$

we may write instead,

$$x = \frac{1}{16} \tan \frac{y}{5},$$

and make our table of values regarding y as the independent variable. The graph is made, regarding y as a function of the independent

y	0	$\frac{5\pi}{6}$	$\frac{5\pi}{3}$	$\frac{5\pi}{2}$	$\frac{10\pi}{3}$	$\frac{25\pi}{6}$	5π
$\frac{y}{5}$	0	$\frac{\pi}{6}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{5\pi}{6}$	π
$\tan \frac{y}{5}$	0	$\frac{1}{3}\sqrt{3}$	$\sqrt{3}$	∞	$-\sqrt{3}$	$-\frac{1}{3}\sqrt{3}$	0
x	0	0.036	0.108	∞	−0.108	−0.036	0

variable x, taking the x scale horizontal and the y scale vertical (Fig. 5). If the character of the tangent function were not known from its geometrical definition, more values would be required in the table in order to construct the graph. The values may be found graphically from the geometrical definition, as in the case of the cosine above, by the construction shown in figure 6. Taking a circle of radius $\frac{1}{16}$, and marking the value of $\frac{y}{5}$ on the circumference, the value of x is found from the intercept of the radius (extended) on the tangent to the circle at the initial line. The scale for x on this tangent line, and for the radius of the circle, are the same as the scale for x in the graph. The graph is discontinuous, x becoming infinite and changing from positive to negative at values of y indicated by the horizontal dotted lines. These dotted lines are **asymptotes** to the curve. A line is an asymptote to a curve when a point on the curve may approach indefinitely close to the line by receding indefinitely far from the origin. The student will easily supply similar treatments for all inverse trigonometric functions.

(c) **Exponential functions.** In this case the independent variable occurs in an exponent. By choosing values of the variable that make an integer of the exponent, a table of values may be constructed by multiplication only. If other values are desired, or if the computations are laborious, use may be made of printed tables of exponential functions.

Consider a rope or belt wound on a drum or pulley, with tensions on the two ends of the rope or belt such that it is just on the point of slipping. We should have $\frac{T_1}{T_2} = e^{\mu\theta}$, where T_1 is the greater tension, T_2 the smaller tension, μ the coefficient of friction, θ the angle subtended at the center of the drum or pulley by that part of rope or belt

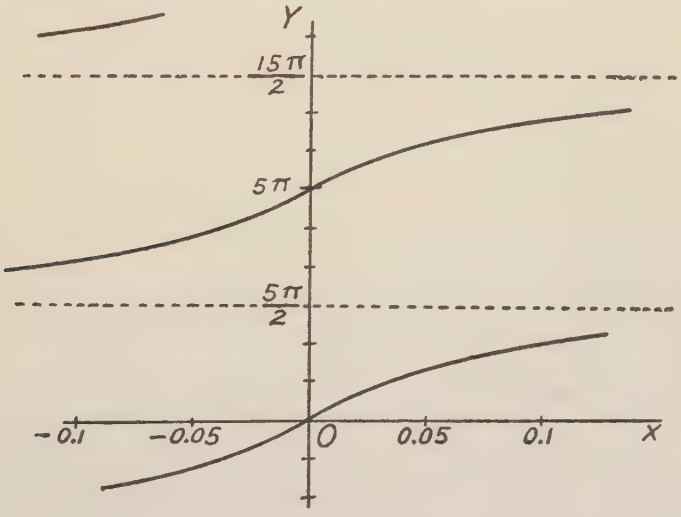


FIG. 5

in contact with it, and e = 2.71828+. This very important number is called the **natural** or **naperian base,** and e is the standard notation for it. We shall find later that exponential functions that employ this base possess properties that are not possessed by those in which any other base is used. Let us suppose $\mu = \frac{1}{3}$, and make a graph showing the ratio $\frac{T_1}{T_2}$ as a function of θ. Taking θ as multiples of three radians, we have integral values of the exponent, and we may compute values of the function. Negative values of the angle are without interest in this problem signifying a reversal of the friction

force, but one negative value is included in the table to help determine the curve near $\theta = 0$ (Fig. 7). In the scale for θ we mark multiples of

θ	-3	0	3	6	9
$\frac{T_1}{T_2}$	0.368	1.0	2.718	7.389	20.086

π, to facilitate the interpretation of θ in degrees, or in laps. The value of $\frac{T_1}{T_2}$ evidently increases very rapidly as θ increases, so that if a rope

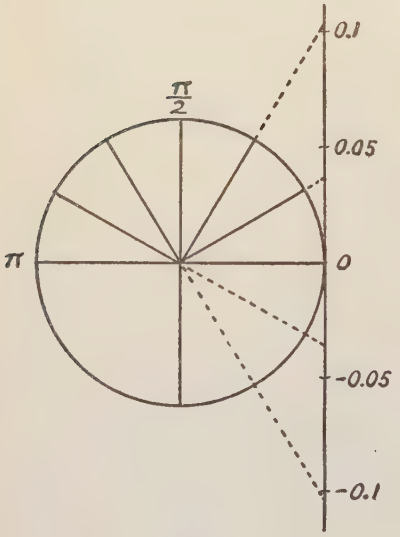


Fig. 6

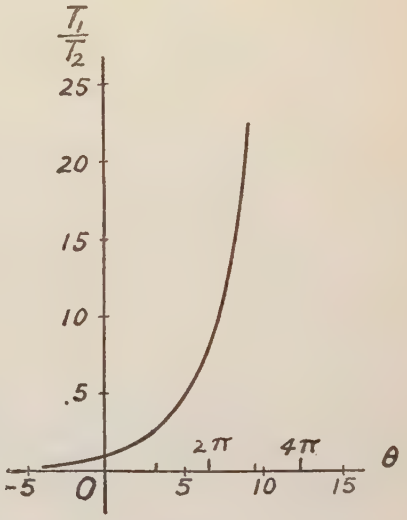


Fig. 7

is given several laps around a horizontal drum a very slight pull at one end will prevent slipping due to a large pull at the other end. For a belt on a pulley the value of θ is usually near π, and the ratio $\frac{T_1}{T_2}$, for $\mu = \frac{1}{3}$, would lie between 2 and 3.

(d) **Logarithmic functions.** These are the inverse of the exponential functions, and may be treated by transforming to exponential functions or by using the values of the logarithms, taken from a table of logarithms.

Given the function

$$y = \log_{10} x^{\frac{1}{5}} = \frac{1}{5} \log_{10} x, \tag{1}$$

we have

$$x = 10^{5y} = e^{\log_e 10 \cdot 5y}. \tag{2}$$

We may see the correctness of this transformation by taking logarithms to the base 10 of all three of the expressions just written,

$$\log_{10} x = 5y \cdot \log_{10} 10 = \log_e 10 \cdot 5y \log_{10} e \tag{3}$$

Now by the definition of logarithm,

$$\log_{10} 10 = 1 \text{ and } \log_e 10 \cdot \log_{10} e = 1,$$

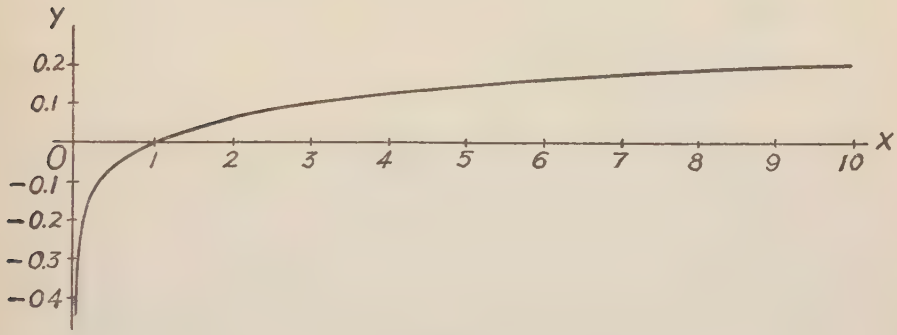


Fig. 8

so that (3) is equivalent to (1). The table of values may be constructed from the equation

$$x = 10^{5y},$$

taking y as multiples of $\frac{1}{5}$. The graph is shown in figure 8. Negative values of y are of equal interest with positive values, but nega-

y.....	$-\frac{3}{5}$	$-\frac{2}{5}$	$-\frac{1}{5}$	0	$\frac{1}{5}$	$\frac{2}{5}$
x.....	0.001	0.01	0.1	1	10	100

tive values of x do not occur. The y-axis is an asymptote to the curve. To make an accurate graph several intermediate values should be used between x = 0.1 and x = 10.

EXERCISES I

1. Show that the volume of a sphere is a rational algebraic function of its radius, but an irrational algebraic function of its surface.

2. Express the diagonal of a rectangle having a constant width of 12 feet as a function of the length of the rectangle. Make a graph of this function.

3. If a rectangle has a constant length of 10 feet, express the angle between the diagonals as a function of the width. Make a graph of this function.

4. Express the temperature in degrees Fahrenheit as a function of the temperature in degrees Centigrade, and make a graph of the function.

5. If $f(x)$ stands for the function $x^2 - 3x + 2$, find $f(0)$, $f(2)$, $f(-3)$, $f(2y)$, $f\left(\frac{x}{2}\right)$.

6. If $\phi(x) = \frac{1-x}{1-2x^2}$, and $u = x^2 + 1$, show that

$$\phi(u) = \frac{x^2}{2x^4 + 4x^2 + 1}.$$

7. If $f(x) = \frac{1}{4}x^2$, show that $f(m + n) - f(m - n) = mn$.

8. If $F(x) = e^{6x}$, show that $\frac{F(u)}{F(v)} = F(u - v)$.

9. If $f(x) = \sin x + a \cos x$, determine a so that $f\left(\frac{\pi}{3}\right) = 0$.

10. The following data were obtained from compression tests of specimens of cast iron. The load, L, is given in thousands of pounds per square inch of section of the specimen, and x, the compression, in thousandths of an inch per inch of length of the specimen.

L.....	8	12	16	20	24	28	32	36	40
x.....	0.5	0.82	1.12	1.43	1.75	2.1	2.48	2.94	3.5
L.....	44	48	49.8						
x.....	4.3	5.4	6						

Make a graph of L as a function of x.

11. Make graphs of the following functions for that part of the x range indicated.

(a) $y = x^2 - \frac{1}{x}$; $x = \frac{1}{10}$ to $x = 4$.

(b) $y = \cos \frac{x}{2}$; $x = 0$ to $x = 4\pi$.

(c) $y = \cot^{-1}x$; $x = -10$ to $x = 10$, least positive y.

(d) $y = 3e^{-10x}$; $x = -0.1$ to $x = 0.4$.

(e) $y = \log_{10}(1 + x)$; $x = 0$ to $x = 10$.

12. For a certain reciprocating engine the position of the piston for certain positions of the crank is stated in the following table. The crank position is given in degrees from dead center position and the position of piston is given in per cent of stroke from the beginning of the out stroke. Make a graph showing the piston displacement as a function of the crank angle. Estimate the speed of the piston at the middle of the stroke, taking the crank radius as one foot, and supposing

Crank angle.....	0	20°	40°	60°	80°	100°	120°
Pist. position.....	0	3.75	14.3	29.7	47.5	64.8	79.7
			140°	160°	180°		
			90.9	97.7	100		

the crank to make 120 revolutions per minute. Assume that the crank turns with uniform angular speed.

Art. 4. Graphs of Compound Functions

We shall understand by a **compound** function a function defined by a combination, in some manner, of two or more simple functions such as we have discussed. It will be convenient to consider the subject in two parts.

(1) **Combination of ordinates.** If the values of a function are to be found by addition, subtraction, multiplication or division of the values of two or more simple functions we may first make graphs of the simple functions, then apply the operations to corresponding ordinates. The scales for independent variables must be the same, as well as the scales for ordinates, in all functions that are to be combined. But it is not necessary that the different functions be evaluated for the same values of the independent variable. Suppose it required to make a graph of the function.

$$y = \sin 2x + 2e^{-\frac{x}{4}} \cos x.$$

We first consider the three simple functions

$$y_1 = \sin 2x, \qquad y_2 = \cos x, \qquad y_3 = 2e^{-\frac{x}{4}}.$$

For the graph of y_1 we construct the table;

x.....	0	$\frac{\pi}{12}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{5\pi}{12}$	$\frac{\pi}{2}$
y_1.....	0	0.5	0.866	1.0	0.866	0.5	0

For the graph of y_2 we have:

x.....	0	$\frac{\pi}{6}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{5\pi}{6}$	π
y_2.....	1	0.866	0.5	0	−0.5	−0.866	−1

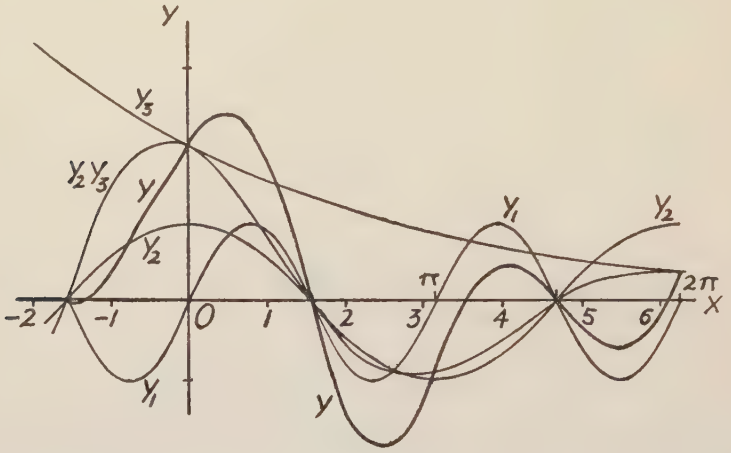


Fig. 9

For y_3, taking convenient values of x:

x.....	−2	0	2	4	6
y_3.....	3.298	2.000	1.214	0.736	0.446

We extend the table for y_3 far enough to include a sufficient portion of the range of x to show a whole wave of y_2 to the right of the origin, and a portion of a wave to the left. A larger portion of the range would be of interest, but the adoption of a smaller scale would obscure some features of the portions shown. Having constructed the graphs of y_1, y_2, and y_3, we construct the product y_2y_3 by multiplying together the ordinates of the separate curves at various values of x. This is easily done at values of x where y_2 or y_3 have values that are integral or simple fractions, as $\frac{1}{2}$, $\frac{1}{4}$, etc. The final step is to add together at various values of x the ordinates of the curves y_1 and y_2y_3, to form the curve y (Fig. 9).

In constructing the tables of values for y_1 and y_2 the values of x used are in terms of π, while for y_3 integral values of x are used. Accordingly, in the scale for x, both integral values and multiples of $\frac{\pi}{2}$ are marked.

The example chosen illustrates the processes of multiplication and addition applied to ordinates of curves. The processes of subtraction and division are equally simple. Addition and subtraction may always be performed graphically in obvious manner. For multiplication and division we may estimate values and multiply numerically, or we may obtain results graphically by a simple device. In the construction shown (Fig. 10), let AB and CD be parallel, then $\frac{OD}{OC} = \frac{OB}{OA}$. If it is desired to find a product y_1y_2, we may take OA = 1, OB =

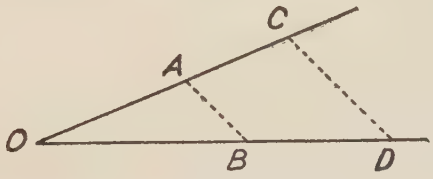


Fig. 10

y_1, OC = y_2, then OD = y_1y_2. If it is desired to find a quotient $\frac{y_1}{y_2}$, we may take OB = 1, OA = y_2, OC = y_1, then OD = $\frac{y_1}{y_2}$.

When all the simple functions involved are such that it is convenient to use the same values of the independent variables in making all the tables of values, the tables may be combined into one, and the values of the compound function may be computed numerically from the tabulated values of the simple functions, thus replacing graphical methods by tabular methods. In making the computations use should be made of slide rule, tables of powers, roots and reciprocals, and any other tabular data available, whenever use of such aids conserves economy of time and accuracy of results.

(2) **A function of a function. Graphical elimination.** We have already dealt with functions which may be regarded as cases of the type considered here. The functions

$$v = \sqrt{4096 - 64h}, \qquad \theta = 10^\circ \cos 4t, \qquad x = 10^{5y},$$

are simple instances, where the dependent variable may be regarded

as a function of a quantity which is itself a function of the independent variable. We may write $v = \sqrt{u}$, or $v^2 = u$, where $u = 4096 - 64h$. The quantity u, so used, is called an **auxiliary** variable, or a **parameter.** Elimination of u between the equations gives the original equation between v and h. We may construct the graph of v as a function of h by first constructing the graphs connecting u and v and u and h. For u and v we have:

v.....	0	±10	±20	±30	±40	±50	60
u.....	0	100	400	900	1600	2500	3600

For u and h the graph is a straight line which may be determined by two pairs of values; h = 0, u = 4096; h = 64, u = 0. The construc-

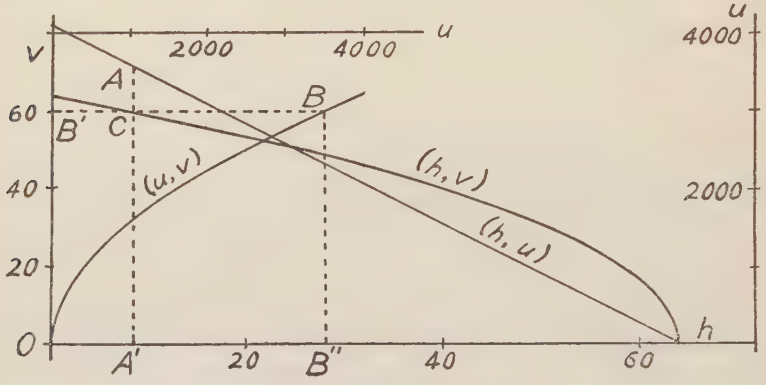


Fig. 11

tion is shown for positive values of v only (Fig. 11) and the three curves are constructed on the same axes to save space. For the (u, v) relation the scale for u is shown at the upper margin, and for the (h, u) relation it is shown at the right hand margin.

Having completed the graph of u as a function of h and that of v as a function of u, we obtain the graph of v as a function of h by a very simple construction. Corresponding to any value of h, as OA′, we have the value A′A of u. Taking OB″ = A′A, we find B″B for the corresponding value of v, then taking A′C = B″B, the point C is a point on the required graph.

In this example, and in case of the other functions mentioned above, the method used here offers no advantages over the simpler methods previously given. There arise cases, however, in which the present method is distinctly advantageous. As a more adequate illustration

consider the function $y = e^{\sin x}$. We treat first the functions $y = e^u$ and $u = \sin x$, choosing scales as before, and obtaining the final graph by the construction used before (Fig. 12). Since sin x is never greater than $+1$ nor less than -1, the range for u is limited, and y is not greater than e, nor less than $\frac{1}{e}$. The function is periodic, the values from $x = 2\pi$ to $x = 4\pi$, from $2n\pi$ to $(2n + 2)\pi$, being the same as from $x = 0$ to $x = 2\pi$.

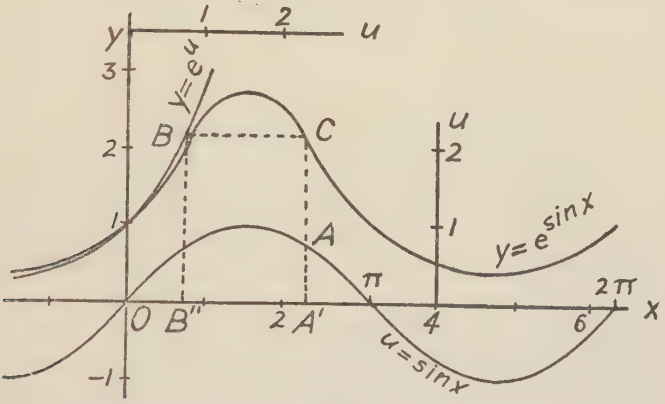


Fig. 12

EXERCISES II

1. Make graphs of the following functions:

(a) $y = 2\cos x + 5\sin x$;

(b) $y = 2\sin\left(x - \frac{\pi}{3}\right) - \frac{1}{2}\sin\left(2x - \frac{\pi}{2}\right)$;

(c) $y = e^{-\frac{1}{5}x}(3\sin x - 2\cos x)$

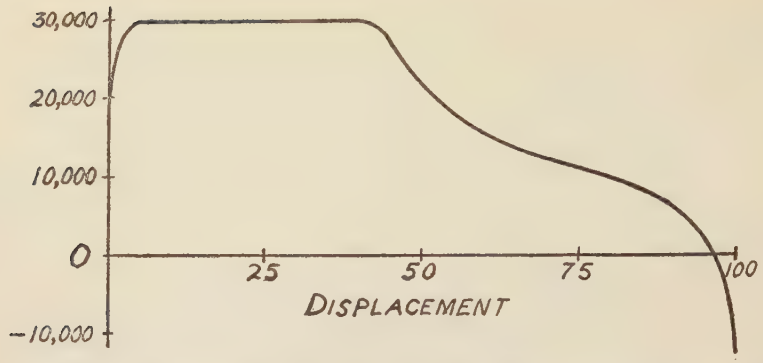
(d) $y = e^{2x} - e^{-2x}$:

(e) $y = 2e^x + 3e^{2x}$:

2. If a particle P resting on a horizontal plane is connected with a point Q by a cord of constant length, a, and if Q moves along a straight line in the plane, the particle P is drawn along a path called a tractrix. If the line on which Q moves is taken as the y-axis, and if Q starts from the origin with the cord perpendicular to the line, the equation is

$$y = a\log\left(\frac{a + \sqrt{a^2 - x^2}}{x}\right) - \sqrt{a^2 - x^2}$$

Construct the curve for the case when $a = 2$.



3. The foregoing graph gives the resultant force F due to steam pressure on the piston of the engine mentioned in exercise 12, I, considered as a function of the piston displacement. The scale for the force F is in pounds, and that for displacement is in per cent of stroke. The graph pertains to the forward stroke only. In exercise 12, I, the displacement is given as a function of the crank angle. From the two functions derive a graph of the force F as a function of the crank angle.

4. In order to find the resultant thrust, P, of the piston rod in the engine mentioned above it is necessary to subtract from F the force, Q, that is required to accelerate the piston and piston rod. An approximate expression for Q as a function of the crank angle θ is

$$Q = 5{,}000(\cos\theta + \frac{1}{4}\cos 2\theta).$$

Using the result of exercise 3, construct a graph of P as a function of the crank angle θ, from 0 to 180°.

5. The function y defined by the equations

$$y = \tan^{-1}u, \qquad u = \frac{1}{2}(e^{x} - e^{-x}),$$

is known as the Gudermannian of x. Construct a graph of the Gudermannian, using u as an auxiliary variable. The standard notation for this function is

$$y = \operatorname{gd} x,$$

and the values are tabulated in most collections of mathematical or physical tables. The student may use any tabulated values as a check for his work.

CHAPTER II

LIMITS AND CONTINUITY

Art. 5. Limits of Functions

It is necessary at this time to present the idea of **limit** as it is used in mathematics. The general idea is contained in the following definition.

If a quantity u, when changing according to a specified law of variation, so behaves that the numerical value of the difference between u and a certain constant k becomes and remains less than any fixed positive number that may be named in advance, then u approaches k as a limit under this law of variation.

For example, the area of a regular polygon of n sides inscribed in a circle approaches the area of the circle as a limit as the number n increases indefinitely. Here the quantity u is the area of the polygon; the law of variation is that the polygon must be regular and inscribed in the given circle, and the number of sides must increase, passing any fixed number; the constant k is the area of the circle. The difference of area, k-u, may be made smaller than any given area by taking n sufficiently large, and moreover, for all larger values of n the difference will remain smaller than the given area.

Now there are other laws of variation for which the area of the polygon will approach the same limit k. The polygon may be circumscribed about the circle instead of inscribed, or we may allow it to be inscribed when n is even, circumscribed when n is odd. In this last case the variable quantity is alternately larger and smaller than its limit, illustrating the important truth that a limit is not necessarily a bound which the variable cannot reach or pass. In its approach to its limit a variable may equal its limit or pass it, but it must ultimately remain nearer to it than any fixed quantity not coincident with the limit.

Before proceeding to special types of limits it is convenient to define certain symbols. The symbols $<$ and $>$ are read "less than" and "greater than" respectively, and they may be combined with the sign of equality as $\leqq$ (less than or equal to) and $\geqq$ (greater than or equal to). These last symbols are sometimes condensed to $\leq$ and $\geq$. The symbol $\neq$ is read "not equal to" or "different from." The expression

$|x|$ is read "absolute value of x" and is the numerical value of x without regard to sign.

Now in any case where u depends on n as in the foregoing example, taking a sequence of values

$$u_1, u_2, u_3, \ldots, u_n, \ldots,$$

if we wish to state that u_n **approaches** k as a limit as n increases indefinitely we may write

$$\lim_{n \to \infty} u_n = k \tag{1}$$

The following condition is necessary and sufficient for the relation (1);

Given any positive number ϵ there exists a number m such that if $n > m$ then $|u_n - k| \leqq \epsilon$.

This condition is **necessary** in the sense that unless it is fulfilled the statement (1) cannot be true; **sufficient** in the sense that if the condition is fulfilled the statement (1) must be true. For a law of variation of the type under consideration the condition is an exact quantitative phrasing of the general definition of limit.

For example, the limit of the sequence

$$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots \ldots \frac{n}{n+1}, \ldots \ldots$$

is 1. Here $u_n = \dfrac{n}{n+1}$ and $|u_n - 1| = \dfrac{1}{n+1}$, so that given any positive ϵ we take $m > \dfrac{1}{\epsilon}$, then for $n > m$ we have

$$|u_n - 1| = \frac{1}{n+1} < \frac{1}{n} < \frac{1}{m} < \epsilon.$$

We now give attention to the notion of limit in relation to a function of a continuous independent variable x. If we wish to say that the function f(x) approaches the limit b as x approaches a we may write

$$\lim_{x \to a} f(x) = b. \tag{2}$$

In applying the general definition here we take $u = f(x)$; the law of variation is contained in the function f(x) and in the fact that x approaches a. The language, or the notation, implies that a is a value in the range of x or a value so related to the range of x that x may

approach it as closely as desired, and further, that by our control over the independent variable x, in causing it to approach a, we may so control the value of the function that it approaches b as a limit.

The necessary and sufficient condition for the limit indicated in (2) may be stated as follows:

Given any positive number ϵ there exists a positive number such that if $|x - a| \leqq \delta$ and if $x \neq a$ then $|f(x) - b| \leqq \epsilon$.

Note that the provision $x \neq a$ excludes from consideration the value of $f(x)$ at $x = a$, namely $f(a)$. This value may or may not be equal to b, but it is not desirable to confuse this question with the question of approach to b as x approaches a.

If the condition for limit is fulfilled and if a sequence of values of x, different from a,

$$x_1, x_2, x_3 \ldots . x_n \ldots . ,$$

is such that x_n approaches a as n increases indefinitely, then the corresponding sequence of function values,

$$f(x_1), f(x_2), \ldots . f(x_n), \ldots . ,$$

will approach b as n increases indefinitely. But the existence of such sequences is not sufficient to imply the limit b, for the statement above pertains to every x different from a such that $|x - a| \leqq \delta$, and not merely to a chosen sequence of such x's. However, it may be shown that the condition for limit is fulfilled if for every such sequence of x's the corresponding sequence of function values approaches b as a limit.

It should be noted further that if the function f(x) is constant, and equal to b for all values of x, then the conditions stated for (2) are fulfilled.

In certain cases it is necessary to consider the behavior of f(x) as x approaches a given value a from one side only, say through values greater than a. To state that f(x) approaches the limit b under this law of variation we may write

$$\lim_{x \to a+} f(x) = b. \tag{3}$$

and similarly, if x is to approach a through values less than a we write

$$\lim_{x \to a-} f(x) = b. \tag{4}$$

Necessary and sufficient conditions for (3) are the same as those stated for (2) except that the provision $x \neq a$ is replaced by $x > a$, while for (4) the provision becomes $x < a$.

Limits of the type (3) and (4) are sometimes called **unilateral** limits, the former the **right handed** limit the latter the **left handed** limit. These terms are justified by graphical considerations.

If the function $f(x)$ approaches b as a limit as x increases indefinitely we may write

$$\lim_{x \to +\infty} f(x) = b. \tag{5}$$

Necessary and sufficient conditions here are:

Given any positive number ϵ there exists a number m such that if $x > m$ then $|f(x) - b| \leqq \epsilon$.

Similarly, if x is to grow indefinitely large in a negative sense we may write.

$$\lim_{x \to -\infty} f(x) = b. \tag{6}$$

Necessary and sufficient conditions for (6) are stated as for (5) except that $x < m$ replaces $x > m$.

In dealing with problems on limits certain fundamental relations that recur frequently should be known and proved once for all. In stating these relations the letters u, v, and y are used to denote functions of the independent variable x, while x is supposed to behave as in some one of the cases (2) to (6). The relations are equally valid if the independent variable is n, as in (1).

If u approaches the limit b and v approaches the limit c then:

(i) **If** $y = u + v$ **then y approaches the limit** $b + c$.

(ii) **If** $y = uv$ **then y approaches the limit** bc.

(iii) **If** $y = \frac{u}{v}$ **and** $c \neq 0$, **then y approaches the limit** $\frac{b}{c}$.

Let us prove (i), assuming that x behaves as in (2). By hypothesis given positive numbers ϵ_1 and ϵ_2, there exist positive numbers δ_1 and δ_2 such that if $|x - a| \leqq \delta_1$ and $x \neq a$ then $|u - b| \leqq \epsilon_1$, and if $|x - a| \leqq \delta_2$ and $x \neq a$ then $|v - c| \leqq \epsilon_2$. We are to show that, given any positive ϵ, there exists a number δ such that if $|x - a| \leqq \delta$ and $x \neq a$ then $|y - (b + c)| \leqq \epsilon$. Now given ϵ let us take $\epsilon_1 = \epsilon_2$

$= \frac{1}{2}\epsilon$ then by taking δ as the smaller of the δ_1 and δ_2 known to be effective, we know that if $|x - a| \leqq \delta$ and $x \neq a$ then

$$|u - b| \leqq \frac{1}{2}\epsilon,$$
$$|v - c| \leqq \frac{1}{2}\epsilon,$$

whence, by adding inequalities and noting that $y = u + v$ we have

$$|y - (b + c)| \leqq \epsilon.$$

If in this proof we had supposed x to behave as in (3) or (4) we should merely have to replace $x \neq a$ by $x > a$ or $x < a$ as the case might be. If x is supposed to behave as in (5) or (6) we are interested in the existence of a number m instead of d. We omit the details in this case, but show the method by considering the proof of (ii) supposing x to increase indefinitely as in (5).

Now, by hypothesis, given any positive ϵ_1 and ϵ_2 there exist m_1 and m_2 such that if $x > m_1$ then $|u - b| \leqq \epsilon_1$ and if $x > m_2$ then $|v - c| \leqq \epsilon_2$. We are to show that, given any positive ϵ there exists m such that if $x > m$ then $|y - bc| \leqq \epsilon$. Let $u = b + u'$ and let $v = c + v'$, then

$$y = uv = bc + bv' + cu' + u'v'.$$

It is convenient to consider three cases.

First when $b = 0$ and $c = 0$. Now we must show that the limit of y is zero. We have $y = u'v'$, and given ϵ we take m_1 and m_2 so that if $x > m_1$ then $|u'| = \epsilon^{\frac{1}{2}}$ and if $x > m_2$ then $|v'| = \epsilon^{\frac{1}{2}}$ then taking m as the greater of m_1 and m_2 we know that for $x > m$ we shall have $|y| \leqq \epsilon$.

Second let one factor approach zero while the other does not, say $b = 0$ and $c \neq 0$. Again we must show that the limit of y is zero. Here $y = cu' + u'v'$. Now u' and v' each approach zero, so by the previous case their product approaches zero. To prove that cu' approaches zero, for an arbitrary ϵ we take m so that for $x > m$ we have $|u'| \leqq \frac{\epsilon}{c}$, then we have clearly $|cu'| \leqq \epsilon$. Now since cu' and $u'v'$ both both approach zero, we see by the relation (i) that y approaches zero.

Third, consider $b \neq 0$ and $c \neq 0$, then $y - bc = bv' + cu' + u'v'$.

The three terms on the right separately approach zero by previous cases, therefore the sum approaches zero by an application of (i) and consequently $y - bc$ approaches zero, from which it follows that y approaches bc.

For the proof of (iii) we show first that the limit of $\frac{1}{v}$ is $\frac{1}{c}$. We have

$$\frac{1}{v} - \frac{1}{c} = \frac{1}{vc}(v - c).$$

Given a positive ϵ, control x so that

$$|v - c| \leqq \frac{\epsilon c^2}{2},$$

while at the same time x is so controlled that

$$\left|v - c\right| \leqq \left|\frac{c}{2}\right|,$$

from which we have $\left|v\right| \geqq \left|\frac{c}{2}\right|$, and therefore

$$\left|\frac{1}{v}\right| \leqq \left|\frac{2}{c}\right|,$$

or

$$\left|\frac{1}{vc}\right| \leqq \left|\frac{2}{c^2}\right|,$$

Then we have

$$\left|\frac{1}{v} - \frac{1}{c}\right| \leqq \left|\frac{1}{vc}\right| \cdot \left|v - c\right| \leqq \frac{\epsilon c^2}{2} \cdot \frac{2}{c^2} \leqq \epsilon.$$

Now $y = u \cdot \frac{1}{v}$, and its limit is $\frac{b}{c}$ by relation (ii).

EXERCISES III

1. Show that $\lim_{x \to 0} \frac{1 + 3x}{2 - 4x} = \frac{1}{2}$, indicating the applications of relations (i), (ii) and (iii) as they occur.

2. Show that $\lim_{x \to \infty} \frac{1 + 3x}{2 - 4x} = -\frac{3}{4}$. (Hint: Divide both numerator and denominator by x.)

3. Show that $\lim_{x \to 0} \sin x = 0$. (Show the way in which the choice of δ depends on the given ϵ).

4. Show that cos x does not approach a limit as x increases indefinitely.

5. Show that $\lim_{x \to -\infty} e^x = 0$. (Show that for a given ϵ an effective m is $m = \log \epsilon$.)

6. Prove that $\lim_{x \to +\infty} e^{-x} \cos x = 0$.

7. Show that $\lim_{x \to \pi} \sec x \cos 2x \sin \frac{1}{2}x = -1$.

8. Show that $\lim_{x \to 0} f(x) = \lim_{x \to 0} f(kx)$ for all values of k and for any function f(x) having a limit as indicated.

9. Given $y = \sin \frac{\pi}{2x}$, show that

(a) If x takes the sequence of values

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots \frac{1}{2n}, \ldots$$

then the corresponding sequence

$$y_1, y_2, \ldots y_n, \ldots$$

has the limit zero.

(b) If x takes the sequence of values

$$1, \frac{1}{5}, \frac{1}{9}, \ldots \frac{1}{4n-3}, \ldots$$

then the corresponding sequence

$$y_1, y_2, \ldots y_n, \ldots$$

has the limit 1.

Using (a) and (b) demonstrate that y does not approach a limit as x approaches zero.

10. Following the theme of exercise 9, show that if $y = f(x)$ and if a sequence of values of x is such that

$$\lim_{n \to \infty} x_n = a$$

and if also the corresponding sequence of values of y is such that

$$\lim_{n \to \infty} y_n = b$$

then if $\lim_{x \to a} f(x)$ exists it must be equal to b.

11. Show that the conditions (3) and (4) considered as simultaneous conditions are equivalent to relation (2).

12. Assuming that $\lim_{x \to o} \cos x = 1$, and that $\lim_{u \to 1} \log u = o$, prove that $\lim_{x \to o} \log \cos x = o$.

13. Following the theme of exercise 12 show that if $u = f(x)$ and $y = F(u)$ and if $\lim_{x \to a} f(x) = k$ and $\lim_{u \to k} F(u) = b$ it follows that $\lim_{x \to a} y = b$.

14. Show that if $f_1(x) \leqq f_2(x)$ for all values of x and if

$$\lim_{x \to a} f_1(x) = a \text{ and } \lim_{x \to a} f_2(x) = b$$

then $a \leqq b$.

Art. 6. Unbounded Variation

In the previous article we have considered the conditions under which a function approaches a limit. In the present article we consider conditions under which a function may fail to approach a limit and yet behave in a perfectly definite manner.

Let us suppose that, as x approaches a, the function f(x) increases in value, passing any fixed number whatever. In this case f(x) is said to **become infinite** as x approaches a. The conventional notation in this situation is

$$\lim_{x \to a} f(x) = +\infty. \tag{1}$$

The necessary and sufficient condition for this relation is as follows:

Given any number h **there exists a positive number** δ **such that if** $|x - a| \leqq \delta$ **and if** $x \neq a$ **then** $f(x) > h$.

It is customary to read equation (1) literally, "the limit of f(x) as x approaches a is plus infinity," but in reality no limit exists. Whatever the language and notation used, the case presented is simply that of a certain mode of **unbounded variation** the only adequate characterization of which is the necessary and sufficient condition just stated.

Analogy with the relations concerning limits discussed in the previous article make it convenient to use the symbol ∞ and the word "infinity," as if they were to represent a number. While no number is thus represented, this symbol and this name may stand for an **ideal element,** not a number, which may fulfil certain relations analogous to algebraic relations and denoted by algebraic symbols. We shall see what these relations are as we proceed.

Other cases of unbounded variation are indicated as follows:

$$\lim_{x \to a} f(x) = -\infty, \tag{2}$$

$$\lim_{x \to a+} f(x) = +\infty, \tag{3}$$

$$\lim_{x \to a-} f(x) = +\infty, \tag{4}$$

$$\lim_{x \to a+} f(x) = -\infty, \tag{5}$$

$$\lim_{x \to a-} f(x) = -\infty, \tag{6}$$

$$\lim_{x \to +\infty} f(x) = +\infty, \tag{7}$$

$$\lim_{x \to -\infty} f(x) = +\infty, \tag{8}$$

$$\lim_{x \to +\infty} f(x) = -\infty, \tag{9}$$

$$\lim_{x \to -\infty} f(x) = -\infty. \tag{10}$$

The student should be able to write the necessary and sufficient conditions for these cases without difficulty. In fact, for (2) the condition is as for (1) except that $f(x) > h$ is changed to $f(x) < h$, and for (3) the condition is as for (1) except that $x \neq a$ is changed to $x > a$. For (7) the condition is:

Given any number h **there exists a number** m **such that if** $x > m$ **then** $y > h$.

Note that here, as also in the previous article where the symbol ∞ was used with the independent variable, the necessary and sufficient conditions are stated entirely in terms of ordinary numbers and algebraic relations, no use being made of the notion "infinity." These conditions may be considered as defining the ideal element ∞; and whatever properties it may possess must grow out of these conditions.

In Art. 5 the fundamental relations (i), (ii) and (iii) are stated for functions u and v which approach limits b and c respectively, with the understanding that b and c are any numbers, including zero, except that in (iii) we have $b \neq 0$. We shall now consider similar relations when u or v or both u and v become infinite. Let it be understood that the independent variable is x and that x behaves as in some one of the cases (1)–(10) above, or (2)–(6) of the previous article.

(i)′ If $y = u + v$ then
(a) If $\lim u = b$ and $\lim v = \pm\infty$ then $\lim y = \pm\infty$,
(b) If $\lim u = +\infty$ and $\lim v = +\infty$ then $\lim y = +\infty$,
(c) If $\lim u = -\infty$ and $\lim v = -\infty$ then $\lim y = -\infty$.

In (a) the limit b of u may be positive, negative or zero. The truth of (a) depends on the fact that given a positive ϵ and an h we may so control x, that $u > b - \epsilon$ and $v > h$, therefore $y > h + b - \epsilon$. By taking h large and ϵ small we can make y larger than any fixed number. It is evident that u and v may be interchanged, in the statement of (a), since $u + v = v + u$.

The student should find it easy to arrange proofs for (b) and (c). Note that in (b) and (c) u, v, and y become infinite in the same sense, all positively or all negatively. If u and v should become infinite in opposite senses, one positive and the other negative, no conclusion can be reached with regard to y without more specific information as to the character of the functions u and v. Considering (i)′ in relation to (i) of the previous article, we find justification for the equations.

$$\text{(a) } a \pm \infty = \pm\infty; \text{ (b) } \infty + \infty = \infty. \tag{11}$$

These are relations, analogous to algebraic relations, fulfilled by the ideal element ∞, but it must not be assumed that they are algebraic equations. For example, if we should transpose in (a) and (b) of (11) we should have $\mp a = \infty - \infty$ and $\infty = \infty - \infty$. These results are contradictory, considered as algebraic equations. The fact is that $\infty - \infty$ is an **indeterminate form,** arising when u and v become infinite in opposite senses as stated above, and a value can be assigned to it only when more specific information is applied.

Corresponding to (ii) of Art. 5 we have

(ii)′ If $y = uv$ then
(a) If $\lim u = b > 0$ and $\lim v = \pm\infty$ then $\lim y = \pm\infty$,
(b) If $\lim u = b < 0$ and $\lim v = \pm\infty$ then $\lim y = \mp\infty$,
(c) If $\lim u = +\infty$ and $\lim v = \pm\infty$ then $\lim y = \pm\infty$,
(d) If $\lim u = -\infty$ and $\lim v = \pm\infty$ then $\lim y = \mp\infty$.

From these propositions, the proofs of which are very simple, we draw the following properties of the ideal element ∞:

$$\text{(a) } (\pm b)(\pm\infty) = \pm\infty; \quad \text{(b) } (\pm\infty)(\pm\infty) = \pm\infty. \tag{12}$$

Here it is understood that the ambiguous sign on the right is + if the two signs on the left are alike, and − if they are different. In (a) of (12) it is understood that $b \neq 0$. If u should approach zero while v becomes infinite the variation of y is not known without more specific information as to the character of u and v. In other words, $0 \cdot \infty$ is an indeterminate form.

It must be remembered that the equations (12) do not signify actual multiplication; they are merely a symbolized statement of (ii)′.

Corresponding to (iii) of Art. 5 we have

(iii)′ If $y = \frac{u}{v}$ then

(a) If $\lim u = b$ and $\lim v = \pm\infty$ then $\lim y = 0$,
(b) If $\lim u = \pm\infty$ and $\lim v = b$ and $v > 0$ then $\lim y = \pm\infty$,
(c) If $\lim u = \pm\infty$ and $\lim v = b$ and $v < 0$ then $\lim y = \mp\infty$,
(d) If $\lim u = \pm b \neq 0$ and $\lim v = 0$ and $v > 0$ then $\lim y = \pm\infty$,
(e) If $\lim u = \pm b \neq 0$ and $\lim v = 0$ and $v < 0$ then $\lim y = \mp\infty$.

In (a) the limit b may be positive, negative or zero. In (b) the limit, b must be zero or positive since v, in its approach to b, is restricted to positive values. Similarly, in (c) v is restricted to negative values and therefore b must be zero or negative. In (d) and (c) u approaches a limit, positive or negative, but not zero, while v approaches 0 through positive values in (d) and negative values in (c).

These propositions lead to the equations

$$\text{(a) } \frac{b}{\infty} = 0; \quad \text{(b) } \frac{\infty}{b} = \infty; \quad \text{(c) } \frac{b}{0} = \infty. \tag{13}$$

If in (iii)′ the limits of u and v are both zero, or if both become infinite, no conclusion can be reached regarding the limit of y without more specific information. We say, therefore, that $\frac{0}{0}$ and $\frac{\infty}{\infty}$ are indeterminate forms.

In these paragraphs we have seen how definite behavior of u and v may lead to definite conclusions, as to the behavior of y. Now under certain conditions conclusions may be drawn when the behavior is not so definite, that is, when one or more of the functions may neither

approach a limit nor become infinite in a definite sense. We state the following propositions without further comment, their validity and usefulness being sufficiently evident.

(i)″ If $y = u + v$ and m is any number, then
(a) If $u > m$ and $\lim v = +\infty$ then $\lim y = +\infty$,
(b) If $u < m$ and $\lim v = -\infty$ then $\lim y = -\infty$.
(ii)″ If $y = uv$, and m is any positive number, then
(a) If $|u| < m$ and $\lim v = 0$ then $\lim y = 0$,
(b) If $u > m$ and $\lim v = \pm\infty$ then $\lim y = \pm\infty$,
(c) If $u < -m$ and $\lim v = \pm\infty$ then $\lim y = \mp\infty$,
(d) If $|u| > m$ and $\lim |v| = +\infty$ then $\lim |y| = +\infty$.
(iii)″ If $y = \frac{u}{v}$, and m is any positive number, then
(a) If $|u| < m$ and $\lim |v| = +\infty$ then $\lim y = 0$,
(b) If $\lim u = \pm\infty$ and $0 < v < m$ then $\lim y = \pm\infty$,
(c) If $\lim u = \pm\infty$ and $-m < v < 0$ then $\lim y = \mp\infty$,
(d) If $\lim |u| = +\infty$ and $|v| < m$ then $\lim |y| = +\infty$,
(e) If $|u| > m$ and $\lim v = 0$ then $\lim |y| = +\infty$.

EXERCISES IV

1. Suggest several functions f(x) for which

$$\lim_{x \to +\infty} f(x) = +\infty.$$

2. Suggest several functions f(x) for which

$$\lim_{x \to 0+} f(x) = +\infty.$$

3. Show that $\lim_{x \to 2+} \frac{x}{x-2} = +\infty$ and $\lim_{x \to 2-} \frac{x}{x-2} = -\infty$.

4. Assign notations to the following:

(a) $\lim_{x \to \frac{\pi}{2}+} \tan x$; (b) $\lim_{x \to \frac{\pi}{2}-} \sec x$; (c) $\lim_{x \to 0+} \cot x$;
(d) $\lim_{x \to +\infty} \text{arc tan } x$; (e) $\lim_{x \to 0-} \csc x$.

5. Write necessary and sufficient conditions for cases (4), (5) and (6) of the present article.

6. Discuss $\lim_{x \to 0+} e^{\frac{1}{x}}$ and $\lim_{x \to 0-} e^{\frac{1}{x}}$, and draw the graph of $y = e^{\frac{1}{x}}$ from $x = -1$ to $x = 1$.

7. (a) Show that $\lim_{x \to 0+} \dfrac{e^{\frac{1}{x}} - e^{-\frac{1}{x}}}{e^{\frac{1}{x}} + e^{-\frac{1}{x}}} = 1$, first dividing both numerator and denominator of the factor by $e^{\frac{1}{x}}$.

(b) Show that $\lim_{x \to 0-} \dfrac{e^{\frac{1}{x}} - e^{-\frac{1}{x}}}{e^{\frac{1}{x}} + e^{-\frac{1}{x}}} = -1$, first multiplying both numerator and denominator of the fraction by $e^{\frac{1}{x}}$.

(c) Make a graph of the function involved in (a) and (b).

8. Write necessary and sufficient conditions for cases (8), (9) and (10) of the present article.

9. Discuss the following:

(a) $\lim_{x \to -\infty} e^x$; (b) $\lim_{x \to +\infty} e^x$; (c) $\lim_{x \to +\infty} \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$; (d) $\lim_{x \to -\infty} \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$.

10. Write proofs of (b) and (c) in (i)′.

11. Discuss the following:

(a) $\lim_{x \to 0+} (x + \cot x)$; (b) $\lim_{x \to 0-} (\cot x + \csc x)$;

(c) $\lim_{x \to \frac{\pi}{2}+} (\tan x + x^2)$; (d) $\lim_{x \to \pi-} (\sec x - \csc x)$.

12. Discuss the following:

(a) $\lim_{x \to 0+} \cot x$; (b) $\lim_{x \to 0} \sin x$; (c) $\lim_{x \to 0+} \sin x \cot x$.

13. Discuss the following:

(a) $\lim_{x \to \frac{\pi}{2}+} \sin x \tan x$; (b) $\lim_{x \to \frac{\pi}{2}-} \sin x \tan x$; (c) $\lim_{x \to 0+} \cot x \log x$;

(d) $\lim_{x \to 0-} e^x \cot x$; (e) $\lim_{x \to -\infty} \operatorname{arc\,sec} x \log x$.

14. Discuss the following:

(a) $\lim_{x \to 1} \dfrac{\sin^{-1} x}{\log x}$; (b) $\lim_{x \to 0+} \dfrac{\log x}{\sin x}$; (c) $\lim_{x \to +\infty} \dfrac{e^x}{\tan^{-1} x}$.

15. Show, by reference to (i)″ that

(a) $\lim_{x\to+\infty}(x + \cos x) = +\infty$; (b) $\lim_{x\to+\infty}(\sec^2 x + e^x) = +\infty$.

16. Show by reference to (ii)″ that

(a) $\lim_{x\to 0} x^2 \sin\frac{1}{x} = 0$; (b) $\lim_{x\to+\infty}(1 - 2\sec^2 x)\log x = -\infty$;

(c) $\lim_{x\to+\infty} |e^x \csc x| = +\infty$.

17. Show by reference to (iii)″ that

(a) $\lim_{x\to\frac{\pi}{2}} \frac{x}{\sec x} = 0$; (b) $\lim_{x\to+\infty} \frac{1 - e^x}{2 - \cos x} = -\infty$.

(c) $\lim_{x\to 0} \left|\frac{\cot x}{\sin\frac{1}{x}}\right| = +\infty$; (d) $\lim_{x\to 0} \left|\frac{1 - \frac{1}{x}}{\log(1 - x)}\right| = +\infty$.

Art. 7. Continuity of Functions

A function f(x) is said to be **continuous** at x = a if the function is defined at x = a, that is, if the definition of the function yields a definite value for f(a), and if also

$$\lim_{x\to a} f(x) = f(a).$$

If either of these two conditions is not fulfilled, the function is said to be discontinuous at x = a. A function is said to be continuous **within the interval from** a **to** b if it is continuous at every value of x between a and b; if also it is continuous at x = a and at x = b then it is said to be continuous **on the interval from** a **to** b **inclusive.**

Discontinuities are of many kinds, but the most important types with which we must deal may be characterized as follows:

First, there may exist a number k such that

$$\lim_{x\to a} f(x) = k,$$

while the function is not defined at a, or if defined then f(a) ≠ k. In this case the function is said to have a **removable** discontinuity at x = a, for we may modify the definition of the function by adding the provision that f(a) = k, then the function becomes continuous at x = a. Illustrations of this type of discontinuity are given by the

functions x csc x and x cot x which take the indeterminate form $0 \cdot \infty$ at $x = 0$. But we clearly have near $x = 0$,

$$|\sin x| < |x| < |\tan x|$$

or, dividing by sin x,

$$1 < |x \csc x| < |\sec x|,$$

and dividing by tan x instead of sin x,

$$|\cos x| < |x \cot x| < 1.$$

Now since $\lim_{x \to 0} \cos x = \lim_{x \to 0} \sec x = 1$, we see that if we assign x csc x and x cot x the value 1 at $x = 0$ they become continuous functions at $x = 0$.

Again there may exist two different numbers, h and k, such that

$$\lim_{x \to a-} f(x) = h, \quad \lim_{x \to a+} f(x) = k.$$

In such case the function may be undefined at a or f(a) may equal h or k or any other number. This is called a **finite** discontinuity. Exercise 7 of the previous article furnishes an example of a finite discontinuity.

Exercise 6 of the previous article furnishes an example of an **infinite** discontinuity. We have,

$$\lim_{x \to 0-} e^{\frac{1}{x}} = 0, \quad \lim_{x \to 0+} e^{\frac{1}{x}} = +\infty,$$

while at $x = 0$ the function is undefined.

A more common type of infinite discontinuity is illustrated by tan x at $x = \frac{\pi}{2}$. Here we have

$$\lim_{x \to \frac{\pi}{2}-} \tan x = +\infty, \quad \lim_{x \to \frac{\pi}{2}+} \tan x = -\infty,$$

while the function is undefined at $\frac{\pi}{2}$. A third illustration is found in the relations

$$\lim_{x \to 1-} \frac{1}{(1-x)^2} = \lim_{x \to 1+} \frac{1}{(1-x)^2} = \lim_{x \to 1} \frac{1}{(1-x)^2} = +\infty.$$

Other types of discontinuity are illustrated by $\sin \frac{1}{x}$, $\tan \frac{1}{x}$, and $\sec \frac{1}{x}$, at $x = 0$. The first of these is continuous everywhere except at $x = 0$, does not become infinite, but fails to have even a unilateral limit at $x = 0$. The second and third not only fail to have unilateral limits at $x = 0$, or even to become infinite in a definite sense, but have infinite discontinuities for every value of x of the form $x = \frac{2}{(2n + 1)\pi}$, where n is any integer, positive or negative.

Since continuity is defined in terms of limits, the investigation of continuity is merely an application of properties of limits stated in Art. 5 and the associated ideas discussed in Art. 6. We see at once that if two functions are continuous their sum, their difference and their product are continuous functions, and their ratio is continuous unless the denominator shall become zero. Furthermore, we see that if one or both of the two functions are discontinuous then the combined function may or may not be discontinuous, and the determination of its character will depend directly on some principle discussed in the two previous articles.

It is worth while to note that if u is a continuous function of x then any **positive integral power** of u is a continuous function of x, and therefore, any **polynomial** in u is a continuous function of x, and the **ratio of two polynomials** is continuous except for those values of x for which the denominator is zero.

A fractional power of u is multiple valued, but will have only one or two real values, and for the present we confine attention to these real values. In case of two real values, one is positive and the other negative, and the two may be considered as different functions for purposes of the present article. With this understanding we may state that a **positive fractional power** of a continuous function is continuous. We shall not give a complete discussion of this proposition, but illustrate the character of the reasoning by use of a particular case. Let $k \neq 0$, and suppose

$$\lim_{x \to a} u = k$$

and let us show that

$$\lim_{x \to a} u^{\frac{1}{3}} = k^{\frac{1}{3}}.$$

We may write the algebraic identity

$$u^{\frac{1}{3}} - k^{\frac{1}{3}} = \frac{u - k}{u^{\frac{2}{3}} + u^{\frac{1}{3}}k^{\frac{1}{3}} + k^{\frac{2}{3}}}$$

As x approaches a the numerator on the right approaches zero, while the denominator does not, therefore the difference on the left must approach zero.

Proof of the continuity of the **elementary transcendental** functions must be based on carefully stated definitions of the functions and on the properties of limits that we have discussed in Art. 5. It is not expedient, however, to give the proof here. We state here certain facts which the student may assume to be true. First, if u is a continuous function of x, the functions e^u, sin u and cos u are a continuous function of x. Second, log u is continuous if u is continuous and positive. Discussion of the continuity of tan u, cot u, sec u and csc u may be based on their expression as rational algebraic functions of sin u and cos u. The inverse trigonometric functions of a quantity u are multiple valued, but if attention is fixed on one of these values, this value is a continuous function of x if u is a continuous function of x and if the value of u is such as to permit a real value for the function in question.

Finally we shall state a very useful general theorem, the proof of which would require methods and principles far beyond the scope of the present work. The student should consider questions of continuity always in association with the graphs of the functions involved. He may satisfy himself as to the plausibility of the following theorem by attempting to draw a graph of a function which shall satisfy its hypothesis but shall violate its conclusions.

Theorem: If a function $y = f(x)$ **is continuous on the interval form** a **to** b **inclusive then there exist two numbers** h **and** k **such that if** $a \leqq x \leqq b$ **then** $h \leqq y \leqq k$, **and such that, given any number** m **such that** $h \leqq m \leqq k$, **there exists a number** c **such that** $a \leqq c \leqq b$ **and** $f(c) = m$.

EXERCISES V

1. Discuss the continuity of the following functions:

(a) $x^2 + \sin x$; (b) $x - \cos x$; (c) $c^x - \sqrt{x^2 + 1}$.

2. Give one or more values of x at which the following functions are discontinuous:

(a) $\dfrac{x + 1}{1 - \cos x}$; (b) $\dfrac{x^2 + 2}{x^2 - 1}$; (c) $\dfrac{\sin 2x + e^x}{(x^2 + 2x + 1)^{\frac{1}{3}}}$.

3. Discuss the continuity of the following:

(a) $e^{\sin x} \cos x$; (b) $(x^2 - 3x + 2)e^{-3x}$; (c) $\dfrac{x + 3 \cos x}{x^2 + x + 1}$.

4. Discuss the discontinuities of the following functions:

(a) $x^3 - 2 \tan x$; (b) $\log \cot x$; (c) $e^x \sec x$;

(d) $\dfrac{1 - \cot x}{1 + \cot x}$; (e) $\dfrac{1 - e^{\frac{1}{x}}}{1 + e^{\frac{1}{x}}}$; (f) $e^{\tan x}$.

5. For each of the following functions state the largest possible interval, containing the value of x indicated, such that the function is continuous within the interval:

(a) $\tan 2\pi x$, $x = 0.7$; (b) $\log \sin x$, $x = 3$;

(c) $\sec x + \cot 2x$, $x = \dfrac{\pi}{6}$; (d) $\dfrac{x^2 - 5}{(x - 1)(x + 2)}$, $x = 0$.

6. Show that the following functions are continuous on the intervals indicated, and in each case determine values of h and k that are effective in the application of the general theorem above.

(a) $3 \sin 2x$, from 0 to π inclusive;

(b) $e^{3x} + x$, from -1 to $+1$ inclusive;

(c) $\log \cos x$, from 0 to $\dfrac{\pi}{3}$ inclusive;

(d) x^2e^{-x}, from -2 to 5 inclusive.

CHAPTER III

DERIVATIVES OF FUNCTIONS

Art. 8. Meaning of Derivatives

In the previous chapter we have seen that if $y = \frac{u}{v}$ and if u and v approach the limits a and b respectively and b is not zero, then y approaches the limit $\frac{a}{b}$. Further, we have seen that if v approaches zero unilaterally while u does not approach zero, then y becomes infinite. If u and v both approach zero we are not able to assign a value to the limit of y without the use of more complete information concerning the functions u and v. In the present chapter we shall be concerned largely with the limits of ratios in which both numerator and denominator approach zero, and we shall be engaged in applying the additional information available in each case to determine the limit of the ratio.

Some of the most familiar ideas involve the limit of such a ratio. We may speak of the speed of a car as it passes a certain crossing. Now what do we mean by speed? Can we time the car for a block and divide the distance by the time to find speed? The car may stop at the middle of the block, and in any case we should find only "average speed." Could we take a shorter distance and divide by the time required to traverse it. We should again have only average speed for that interval of time. Let s denote distance from the point and t the time required to traverse this distance, then $\frac{s}{t}$ is the average speed during the time t. Now by the speed at the point we mean the limit that $\frac{s}{t}$ approaches as t approaches zero. We may express this speed in feet per second or miles per hour, etc., according to the units in which s and t are measured, but this does not mean that the car goes any number of feet in a given time, or that it goes any part of a mile in a given time.

Again we speak of the intensity of pressure at a point on a surface under water or other liquid. What do we mean? If we take the resultant force F on a portion of the surface and divide the force by the area A of the portion of surface, we obtain "average intensity."

By intensity of pressure at a point we mean the limit $\frac{F}{A}$ as the area A approaches zero, always containing the point in question.

Now as this question presents itself in problems requiring mathematical treatment, it always pertains to a function of a continuous variable. Let us, for present purposes, call the variable x, and let the function be represented by

$$y = f(x). \tag{1}$$

Let y_o be the value of the function corresponding to the value x_o of x, i.e.,

$$y_o = f(x_o),$$

and let us suppose that the function is continuous, so that as x approaches x_o the function value y approaches y_o. Here $y - y_o$ approaches zero as $x - x_o$ approaches zero, and we are concerned with the limit of the ratio $(y - y_o)/(x - x_o)$ as $x - x_o$ approaches zero. It is customary to denote $y - y_o$ by Δy, and $x - x_o$ by Δx, the symbol Δ, the Greek delta, denoting "difference." We call these quantities the y-difference and the x-difference, or the **increment** in y, the **increment** in x.

We have $x = x_o + \Delta x$, $y = y_o + \Delta y$,
so that (1) becomes,

$$y_o + \Delta y = f(x_o + \Delta x),$$

But

$$y_o = f(x_o),$$

therefore

$$\Delta y = f(x_o + \Delta x) - f(x_o).$$

Now dividing by Δx,

$$\frac{\Delta y}{\Delta x} = \frac{f(x_o + \Delta x) - f(x_o)}{\Delta x}.$$

The limit of the left member of the equation is equal to the limit of the right member; in symbols,

$$\lim_{\Delta x \to o} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to o} \frac{f(x_o + \Delta x) - f(x_o)}{\Delta}. \tag{2}$$

In many cases the right hand member fo this equation can be found very directly, and in any case it affords our only method of approaching the problem in a mathematical sense. The value of this limit is called the **derivative of y with respect to x at x_0.**

Now let us apply this idea to the question of speed. If an object is moving, then its distance s from a certain point is a function of the time t. To be specific, suppose

$$s = 16t^2 - 10t,$$

where s is in feet and t is in seconds. When $t = 1$, $s = 6$, and it can be seen that the function is continuous at $t = 1$. To find the speed when $t = 1$, we have, putting $t = 1 + \Delta t$, $s = 6 + \Delta s$,

$$6 + \Delta s = 16(1 + \Delta t)^2 - 10(1 + \Delta t) = 6 + 22\,\Delta t + 16\,(\Delta t)^2$$

or

$$\Delta s = 22\Delta t + 16(\Delta t)^2.$$

Dividing by Δt we have

$$\frac{\Delta s}{\Delta t} = 22 + 16\Delta t.$$

This equation shows the ratio between the distance and the time interval. If $\Delta t = 1$, $\Delta s = 38$, showing that from $t = 1$ to $t = 2$ the object will move 38 feet. If $\Delta t = \frac{1}{2}$, $\Delta s = 15$, showing that from $t = 1$ to $t = \frac{3}{2}$, the object will move 15 feet, or that the average speed during the half second is 30 feet per second. If $\Delta t = -\frac{1}{2}$, $\Delta s = -7$, showing that the average speed during the previous half second is 14 feet per second. It is quite plain that,

$$\lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = 22,$$

and that the speed at the moment $t = 1$ is 22 feet per second.

If a function

$$y = f(x)$$

is continuous for every value of x in a certain range, and if the limit which we call the derivative of y with respect to x can be found at each value of x in the range, then this derivative is a new function of x, having a definite value for each value of x on the range. We may in this case find the limit of $\frac{\Delta y}{\Delta x}$ in terms of x, without specifying the place x_0 from which Δx is measured. This new function is called the **derived function,** or the **derivative with respect to** x **of** f(x). It may have a separate notation of its own, as $\phi(x)$, $f_1(x)$, etc., but when it is desired to indicate its character as a derived function we usually adopt one of the symbols

$$f'(x), \quad \frac{d}{dx} f(x), \quad \frac{dy}{dx}, \quad D_x y, \quad D_x f(x).$$

The symbols $\frac{d}{dx}$ and D_x may be regarded as abbreviations for the words "derivative with respect to x," and when followed by y or f(x) they indicate that the process of finding the derivative called **differentiation,** has been applied to the function to derive a new function. We may regard as the definition of a derived function, or a derivative, the following equation:

$$\frac{dy}{dx} = \frac{d}{dx} f(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \tag{3}$$

Applying this to the function

$$s = 16t^2 - 10t$$

we have

$$\frac{f(t+\Delta t) - f(t)}{\Delta t} = \frac{[16(t+\Delta t)^2 - 10(t+\Delta t)] - [16t^2 - 10t]}{\Delta t}.$$

The second member reduces to

$$32t - 10 + 16\Delta t,$$

so that

$$\frac{ds}{dt} = \frac{d}{dt}(16t^2 - 10t) = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = 32t - 10.$$

This is a new function of t, the derived function, and its value for any t is the speed at that time. When t = 1, the speed is 22, as before,

but when t = 2, the speed is 54, and so at each moment a definite speed is given.

Suppose the curve in Figure 13 to be the graph of the function y = f(x). Let P′P be the value of y corresponding to the value OP′ of x. If we wish to find the derivative for this value of x we consider the increment $\Delta x = P'Q'$, and the corresponding increment $\Delta y = RQ$ in y. As Δx approaches zero, the secant line PQ swings about P, and approaches the position PA. The triangle PQR becomes smaller and smaller, but its shape approaches the shape of PAB, and the ratio $\frac{\Delta y}{\Delta x}$ approaches $\frac{BA}{PB}$. The line PA is the tangent line to the graph at P, and the relations presented here enable us to estimate the value

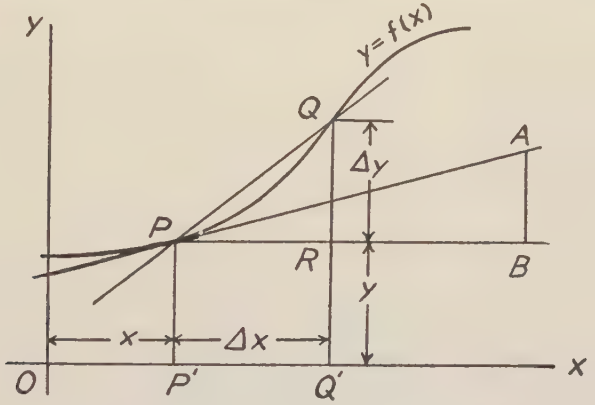


Fig. 13

of a derivative from the graph of a function by drawing a tangent line, as PA and a horizontal line as PB, and estimating the value of the vertical BA measured in the scale of y and of the horizontal PB measured in the scale of x.

As x varies, passing through the value OP′, y varies passing through the value P′P, and it is seen that, at that instant, the rate of increase of y is to the rate of increase of x as BA is to PB. The derivative of one variable with respect to another thus furnishes a mode of comparing their instantaneous rates of increase. In fact, if we wish to state in words the meaning of derivative we can do no better than to say that **the derivative of a function of x with respect to x is the instantaneous rate of change of the value of the function measured in terms of the rate of change of x at the same instant.**

Since x is an independent variable, we assign it values at will, and may regard it as increasing or decreasing at any rate we choose. The function changes with x, and its rate of change is dependent on that of x, and becomes a definite thing only when measured in terms of that of x. The speed of an object is the rate of increase of distance measured in terms of the rate of increase in time. If we say that the speed at a certain instant is 10 feet per second, we mean that the distance, measured in feet, is increasing 10 times as fast as the time, measured in seconds.

EXERCISES VI

1. Find the derivative of the function

$$y = 2x^2 - 6x + 5,$$

at the place $x = 2$, first by considering the limit of the ratio of the increments $\Delta y = y - 1$ and $\Delta x = x - 2$, and second, by finding the derived function and taking its value at $x = 2$.

2. A particle moves in such manner that the distance s in feet at the end of t seconds is given by the equation.

$$s = 40 - 2t - 8t^2.$$

Find the position and speed of the particle at the end of 2 seconds; at the end of one second; at the beginning of the count of time. Considering acceleration as the rate of change of speed, find the acceleration at each of the three instants mentioned.

3. The area of a square is a function of the length of a side. Express this function, find the derived function, and thus show that the rate of increase of area measured in terms of the rate of increase of the side, is equal to half the perimeter of the square.

4. Express the volume of a cube as a function of the length of an edge, and show the value of the derived function is half the area of the surface of the cube.

5. Show that the rate of increase of area of a circle is equal to the circumference times the rate of increase of the radius.

6. Express the volume of a sphere as a function of the radius, and show that the derived function is an expression for the surface of the sphere as a function of the radius, and that the derivative of this function (the second derivative or second derived function) is equal to four times the circumference of the sphere.

7. Show that the derivative of the area of a rectangle with respect to the length is equal to the breadth, and that the derivative of the

volume of a right prism with respect to the altitude is equal to the area of the base.

8. Suppose that the graph of a certain function $y = f(x)$, is the arc of a circle with center at the origin, and that the scales for x and y are equal. Show, by geometrical means, that at any value of x the value of the derivative is given by

$$\frac{dy}{dx} = -\frac{x}{y}.$$

9. From a graph of $y = \log_{10} x$ estimate the value of $\frac{dy}{dx}$ at $x = 1$; $x = 2$; $x = 10$.

Art. 9. General Rules for Finding Derivatives

Using the definition of derivative, or derived function, given in the last article, we now derive a few rules that enable us to obtain derivatives without returning to a consideration of limits in each case. We adopt the letters y, u, v, to denote functions of the independent variable x, and we usually consider u and v, as auxiliary variables, used to express y as a function of x. We state the rules here, then prove them, and show their uses later.

Rule I. If $y = u + v$, then $\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}$.

Rule II. If $y = uv$, then $\frac{dy}{dx} = v\frac{du}{dx} + u\frac{dv}{dx}$.

Rule III. If $y = u^n$, then $\frac{dy}{dx} = nu^{n-1}\frac{du}{dx}$.

Rule IV. If $y = \frac{u}{v}$, then $\frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$.

Rule V. If $y = f(u)$, then $\frac{dy}{dx} = \frac{d}{du}f(u) \cdot \frac{du}{dx}$.

Proof of Rule I. Here y is the sum of two functions of x, namely u and v. Let us denote by the symbols Δu, Δv, and Δy the increments in these variables that correspond to the increment Δx in the independent variable. We have then,

$$y + \Delta y = u + \Delta u + v + \Delta v,$$

but $y = u + v$, therefore, subtracting and dividing through by Δx, we have

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}.$$

Now as Δx approaches zero, these three "difference quotients" approach values called derivatives, and these derivatives fulfill the equation below, by proposition (i) of Art. 5:

$$\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}.$$

Proof of Rule II. In this case we have $y = uv$ and

$$y + \Delta y = (u + \Delta u)(v + \Delta v) = uv + v\,\Delta u + (u + \Delta u)\,\Delta v$$

therefore, subtracting and dividing by Δx

$$\frac{\Delta y}{\Delta x} = v\frac{\Delta u}{\Delta x} + (u + \Delta u)\frac{\Delta v}{\Delta x}.$$

As Δx approaches zero the three difference quotients approach the derivatives, and $u + \Delta y$ approaches u, so that, using the proposition (ii) of Art. 5 we have

$$\frac{dy}{dx} = v\frac{du}{dx} + u\frac{dv}{dx}.$$

Proof of Rule III. When n is a positive integer: First let us consider y as the product of several factors, say n factors.

$$y = u_1u_2u_3 \ldots u_n.$$

Applying Rule II we have

$$\frac{dy}{dx} = u_2u_3 \ldots u_n\frac{du_1}{dx} + u_1\frac{d}{dx}(u_2u_3 \ldots u_n)$$

$$= u_2u_3 \ldots u_n\frac{du_1}{dx} + u_1u_3 \ldots u_n\frac{du_2}{dx} + u_1u_2\frac{d}{dx}(u_3 \ldots u_n)$$

$$= u_2u_3 \ldots u_n\frac{du_1}{dx} + u_1u_3 \ldots u_n\frac{du_2}{dx} + \ldots + u_1 \ldots u_{n-1}\frac{du_n}{dx},$$

there being n terms, all the terms that can be formed by taking the derivative of one factor and multiplying by all the other factors. Now if all these factors are the same function, $u_1 = u_2 = \ldots = u_n = u$, then

$$y = u^n$$

and

$$\frac{dy}{dx} = nu^{n-1}\frac{du}{dx}.$$

When n is a positive fraction: Let $n = \frac{p}{q}$ where p and q are positive integers, then

$$y = u^n = u^{\frac{p}{q}} \text{ or } y^q = u^p.$$

Now applying the rule for the case already proved, we have

$$qy^{q-1}\frac{dy}{dx} = pu^{p-1}\frac{du}{dx},$$

or

$$\frac{dy}{dx} = \frac{p}{q}\frac{u^{p-1}}{y^{q-1}}\frac{du}{dx}.$$

But

$$y^{q-1} = u^{\frac{p}{q}(q-1)} = u^{p-\frac{p}{q}},$$

and

$$\frac{u^{p-1}}{y^{q-1}} = \frac{u^{p-1}}{u^{p-\frac{p}{q}}} = u^{\frac{p}{q}-1} = u^{n-1},$$

so that we have

$$\frac{dy}{dx} = nu^{n-1}\frac{du}{dx}.$$

When n is negative: Let $n = -m$, where m is a positive fraction or integer, then

$$y = u^n = u^{-m} = \frac{1}{u^m}$$

or, multiplying by u^m,

$$yu^m = 1.$$

Applying Rule II, we have

$$y \frac{d}{dx}(u^m) + u^m \frac{dy}{dx} = \frac{d}{dx}(1).$$

Now when we consider the derivative as a rate of change, we see that the derivative of a constant is zero, and applying the cases already treated for Rule III, we have

$$\frac{d}{dx}(u^m) = mu^{m-1} \frac{du}{dx},$$

therefore

$$myu^{m-1} \frac{du}{dx} + u^m \frac{dy}{dx} = 0$$

or, solving for $\frac{dy}{dx}$,

$$\frac{dy}{dx} = -myu^{-1} \frac{du}{dx} = -mu^{-m-1} \frac{du}{dx} = nu^{n-1} \frac{du}{dx}.$$

Proof of Rule IV. We may take,

$$y = \frac{u}{v} = uv^{-1},$$

then applying Rules II and III, we have

$$\frac{dy}{dx} = v^{-1} \frac{du}{dx} + u \frac{d}{dx}(v^{-1}) = v^{-1} \frac{du}{dx} - uv^{-2} \frac{dv}{dx} = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}.$$

Proof of Rule V. Here we have the two equations

$$y = f(u)$$

$$y + \Delta y = f(u + \Delta u).$$

Subtracting, dividing by Δx, and multiplying the second member by Δu in both numerator and denominator,

$$\frac{\Delta y}{\Delta x} = \frac{f(u + \Delta u) - f(u)}{\Delta u} \cdot \frac{\Delta u}{\Delta x}.$$

Now as Δx approaches zero the difference quotients approach derivatives, and we have, applying proposition (ii) of Art. 5,

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

Art. 10. Derivatives of Algebraic Functions

Besides the four general rules proved in Art. 6, we have use for the two special rules:

Rule VI. The derivative of a constant is zero.

Rule VII. The derivative of a variable with respect to itself is unity, i.e., $\frac{dx}{dx} = 1$.

The proofs of these rules are left to the student.

(1) **Polynomials.** A polynomial is the sum of products of a constant by a power of the variable. By Rule I we differentiate each term separately and add results. Each term being of the form cx^n, where c is a constant, we have by Rules II and III,

$$\frac{d}{dx} cx^n = ncx^{n-1} \frac{dx}{dx} + x^n \frac{dc}{dx} = ncx^{n-1},$$

the last reduction resulting from Rules VI and VII. As an example we may take

$$y = 6x^2 - 3x^3 + \frac{1}{3} x^5 - 8,$$

and the derivative is

$$\frac{dy}{dx} = 12x - 9x^2 + \frac{5}{3} x^4.$$

(2) **Rational algebraic functions.** These may involve quotients, but they may always be written as sums of products, the denominators becoming factors with negative exponents. We illustrate by example:

$$y = \frac{(x - 2)(x^2 + 3)}{x^2 - 2x + 4} - \frac{1}{(x - 3)^2}.$$

First applying Rule I we differentiate the two terms separately. For the derivative of the first term we have by Rule IV,

$$\frac{(x^2 - 2x + 4)\frac{d}{dx}[(x - 2)(x^2 + 3)] - (x - 2)(x^2 + 3)\frac{d}{dx}(x^2 - 2x + 4)}{(x^2 - 2x + 4)^2}.$$

Now

$$\frac{d}{dx}(x - 2)(x^2 + 3) = (x - 2)\frac{d}{dx}(x^2 + 3) + (x^2 + 3)\frac{d}{dx}(x - 2)$$

$$= (x - 2)(2x) + (x^2 + 3)(1) = 3x^2 - 4x + 3$$

while

$$\frac{d}{dx}(x^2 - 2x + 4) = 2x - 2.$$

Substituting these values and reducing we have

$$\frac{x^4 - 4x^3 + 13x^2 - 4x}{(x^2 - 2x + 4)^2}.$$

Now the second term is treated by Rule III,

$$\frac{d}{dx}(x - 3)^{-2} = -2(x - 3)^{-3} = \frac{-2}{(x - 3)^3},$$

so that the final result is,

$$\frac{dy}{dx} = \frac{x^4 - 4x^3 + 13x^2 - 4x}{(x^2 - 2x + 4)^2} + \frac{2}{(x - 3)^3}.$$

(3) **Irrational algebraic functions.** These functions involve the roots of algebraic expressions, which may be treated by the adoption of fractional exponents. For example let

$$y = \frac{\sqrt[3]{x^2 - 2}}{x^2 - \sqrt{4 - 6x}}.$$

This may be written

$$y = (x^2 - 2)^{\frac{1}{3}}[x^2 - (4 - 6x)^{\frac{1}{2}}]^{-1}$$

and applying Rules II and III, then I, VI, and VII

$$\frac{dy}{dx} = [x^2 - (4 - 6x)^{\frac{1}{2}}]^{-1} \cdot \frac{1}{3}(x^2 - 2)^{-\frac{2}{3}} \frac{d}{dx}(x^2 - 2)$$

$$+ (x^2 - 2)^{\frac{1}{3}} (-1)[x^2 - (4 - 6x)^{\frac{1}{2}}]^{-2} \frac{d}{dx}[x^2 - (4 - 6x)^{\frac{1}{2}}] \cdot$$

$$= \frac{2x}{3}[x^2 - (4 - 6x)^{\frac{1}{2}}]^{-1}(x^2 - 2)^{-\frac{2}{3}}$$

$$- (x^2 - 2)^{\frac{1}{3}}[x^2 - (4 - 6x)^{\frac{1}{2}}]^{-2}[2x + 3(4 - 6x)^{-\frac{1}{2}}].$$

This may be written without fractional or negative exponents,

$$\frac{dy}{dx} = \frac{\frac{2x}{3}(x^2 - \sqrt{4 - 6x}) - (x^2 - 2)\left(2x + \frac{3}{\sqrt{4 - 6x}}\right)}{\sqrt[3]{(x^2 - 2)^2}\,(x^2 - \sqrt{4 - 6x})^2}.$$

EXERCISES VII

1. Find the derivatives of the following functions.

(a) $3x - 6x^2$;
(b) $x^3 - 2x^2 + 7x - 6$;
(c) $(1 - x)^2 + 4(x^2 - 7x + 2)$;
(d) $(1 - x^2)^2 + (x^2 - 5)^3$;
(e) $(x^3 + 3x - 2)^5$;
(f) $x^4 - \frac{2}{3}x^2 + (2x - 3)^2$.

2. Show that if $y = \frac{c}{u^n}$, where c is a constant and n is a positive fraction or integer, and u is any function of x,

$$\frac{dy}{dx} = -\frac{nc}{u^{n+1}} \frac{du}{dx}.$$

3. Find $\frac{dy}{dx}$ in each of the following cases:

(a) $y = (3 - x^2)(x - 4x^2 + 2)$;
(b) $y = x^2(x^2 - 3)^3$;
(c) $y = 10x^3(16 - 4x^2)^2$;
(d) $y = (x^2 - 6x + 2)^2(x^4 - 1)$.

4. Use the result stated in exercise 2 above to find derivatives of the following functions with respect to the variable indicated in each case.

(a) $\frac{6}{1 - x^2}$, x;
(b) $\frac{1}{y^2 + 2y - 5}$, y;
(c) $\frac{a}{(a - 2t)^n}$, t;
(d) $\frac{c}{a + bs^k}$, s.

5. (a) If $y = \dfrac{x-2}{x-3}$, find $\dfrac{dy}{dx}$; (c) If $v = \dfrac{u^2}{u^3-1}$, find $\dfrac{dv}{du}$;

(b) If $F = \dfrac{k}{r^2}$, find $\dfrac{dF}{dr}$; (d) If $s = 6t - \dfrac{1}{(t-1)^2}$ find $\dfrac{ds}{dt}$.

6. Find the following derived functions:

(a) $\dfrac{d}{dx}(x^{\frac{1}{2}} - 6x^{-1})$; (d) $\dfrac{d}{ds}\dfrac{(s-2)^2}{(s-1)^{\frac{3}{2}}}$;

(b) $\dfrac{d}{dt}(t + \sqrt{t^2-4})$; (e) $\dfrac{d}{dr}\dfrac{a(1-r^n)}{1-r}$;

(c) $\dfrac{d}{du}\dfrac{3u}{\sqrt{1-u^2}}$; (f) $\dfrac{d}{dp}\dfrac{p}{\sqrt{(q^2-p^2)^2 + b^2p^2}}$.

7. If r, s and v denote respectively the radius, surface and volume of a sphere, in inches, square inches and cubic inches, find $\dfrac{ds}{dr}$, $\dfrac{dv}{dr}$ and $\dfrac{dv}{ds}$. Show that $\dfrac{dv}{ds} = \dfrac{dv}{dr} \Big/ \dfrac{ds}{dr}$, and show that the relation is a special case of a general rule that follows from Rule V, provided that the function u is not constant in Rule V.

8. The "law of falling bodies" is expressed by the equation

$$s = \frac{1}{2}gt^2 + v_0t + s_0,$$

where s is distance fallen, t is time and v_0 and s_0 are initial speed and initial distance, respectively, and g is constant. Find the speed when $t = 5$; when $t = 8$. Find the rate of change of speed, i.e., the acceleration, when $t = 5$; when $t = 8$.

9. A conical vessel, with circular base two feet in diameter and three feet high, is being filled with water through an orifice at the vertex. If water is flowing in at the rate of 100 cu. in. per minute, how fast is the surface rising when the water is six inches deep? When 18 inches deep?

10. A ladder 10 feet long rests against a vertical wall, and its foot slips out from the wall on a horizontal floor at the rate of two feet per second. How fast is the top of the ladder slipping down when the foot is four feet from the wall? When six feet from the wall?

11. If a projectile is fired at an angle of elevation α, with a muzzle velocity v_0, the horizontal speed is constantly $v_0 \cos \alpha$, and the height,

y, of the projectile is given as a function of the horizontal distance, x, by means of the equation

$$y = x \tan \alpha - \frac{gx^2}{2v_0^2 \cos^2 \alpha}.$$

Find the rate at which the height is increasing, as a function of x. For what value of x will the height stop increasing, and how high will the projectile be at that moment? In what direction will the projectile be moving at that moment? What will be the direction of motion when $x = 10v_0$?

12. Two ships start from the same port at the same time. One ship sails east at 12 knots an hour, the other north at 16 knots an hour. How fast are the ships separating at the end of 15 minutes? At the end of an hour? If the ship sailing east starts 30 minutes before the other, how fast are they separating 15 minutes after the second ship starts?

Art. 11. Derivatives of Trigonometric and Inverse Trigonometric Functions

We require here a few additional special rules, all of which are stated at this point, and the proofs are discussed later.

Rule VIII. (a) $\frac{d}{dx} \sin x = \cos x;$

(b) $\frac{d}{dx} \cos x = -\sin x;$

(c) $\frac{d}{dx} \tan x = \sec^2 x;$

(d) $\frac{d}{dx} \cot x = -\csc^2 x;$

(e) $\frac{d}{dx} \sec x = \sec x \tan x;$

(f) $\frac{d}{dx} \csc x = -\csc x \cot x.$

Rule IX. (a) $\frac{d}{dx} \sin^{-1} x = \pm \frac{1}{\sqrt{1 - x^2}} \begin{bmatrix} +, \text{ 1st \& 4th quadrants} \\ -, \text{ 2nd \& 3rd } \text{“} \end{bmatrix};$

(b) $\frac{d}{dx} \cos^{-1} x = \mp \frac{1}{\sqrt{1 - x^2}} \begin{bmatrix} -, \text{ 1st \& 2nd quadrants} \\ +, \text{ 3rd \& 4th } \text{“} \end{bmatrix};$

(c) $\frac{d}{dx} \tan^{-1} x = \frac{1}{1 + x^2};$

$$\text{(d)}\ \frac{d}{dx}\cot^{-1}x = \frac{-1}{1+x^2};$$

$$\text{(e)}\ \frac{d}{dx}\sec^{-1}x = \pm\frac{1}{x\sqrt{x^2-1}}\begin{bmatrix}+, \text{1st \& 2nd quadrants}\\ -, \text{3rd \& 4th}\quad ``\end{bmatrix};$$

$$\text{(f)}\ \frac{d}{dx}\csc^{-1}x = \mp\frac{1}{x\sqrt{x^2-1}}\begin{bmatrix}-, \text{1st \& 4th quadrants}\\ +, \text{2nd \& 3rd}\quad ``\end{bmatrix}.$$

Before taking up the proof of these rules we prove the proposition: **On any given circle, the ratio of the length of a chord to that of the arc it subtends approaches unity as the length of the arc approaches zero.**

Let PQ be the chord subtending the arc PLQ of the circle (Fig. 14) whose center is at 0. We wish to prove that

$$\lim_{PLQ\to 0}\frac{PQ}{PLQ} = 1.$$

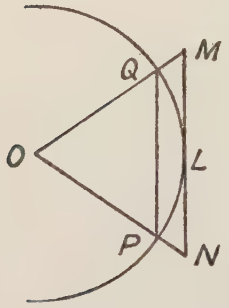


Fig. 14

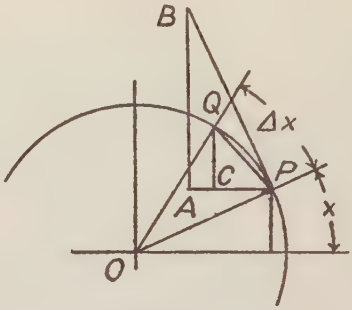


Fig. 15

Proof: Take MN parallel to PQ and tangent to the arc. By similar triangles we have $\frac{PQ}{MN} = \frac{OP}{OM}$, but as the length of arc PLQ approaches zero OM approaches the constant OP, therefore $\frac{PQ}{MN}$ approaches unity. Now PQ < PLQ < MN, therefore

$$1 > \frac{PQ}{PLQ} > \frac{PQ}{MN}$$

As the arc approaches zero, this last ratio approaches 1, and the quantity that lies between it and 1 must therefore approach 1.

Proof of Rule VIII (a). Let us take a circle of unit radius (Fig. 15)

then the measure in radians of the angle x subtended by the arc, TP is the length of the arc, and the increment Δx, in radians, is equal to the length of arc PQ. Also $\sin x = \frac{RP}{OP} = RP$, and the increment in the sine corresponding to the increment Δx is CQ. Therefore, if $y = \sin x$, we have

$$\frac{\Delta y}{\Delta x} = \frac{CQ}{\text{arc } PQ} = \frac{CQ}{\text{chord } PQ} \cdot \frac{\text{chord } PQ}{\text{arc } PQ}.$$

Now as Δx approaches zero, the direction of the chord PQ approaches the direction of the tangent PB, and the ratio $\frac{CQ}{\text{chord } PQ}$ approaches the ratio $\frac{AB}{PB}$, but this is $\cos x$, since the angle ABP is equal to x; and since the ratio of the chord to the arc approaches 1, we have

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \cos x$$

The student will readily see that this proof is valid for any value of x, and should draw figures to show the relations in other quadrants.

Proof of Rule VIII (b). Using the same figure, we have

$$y = \cos x = \frac{OR}{OP} = OR,$$

and

$$\frac{\Delta y}{\Delta x} = \frac{-CP}{\text{arc } PQ} = \frac{-CP}{\text{chord } PQ} \cdot \frac{\text{chord } PQ}{\text{arc } PQ},$$

from which, reasoning as before, we have

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\frac{AP}{PB} = -\sin x.$$

Proof of Rule VIII (c). We take

$$y = \tan x = \frac{\sin x}{\cos x},$$

then applying Rules IV we have for the derivative

$$\frac{dy}{dx} = \frac{\cos x \frac{d}{dx}(\sin x) - \sin x \frac{d}{dx}(\cos x)}{\cos^2 x}$$

and performing the differentiations indicated, by the rules just proved,

$$\frac{dy}{dx} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \sec^2 x.$$

The proofs of the remaining parts of Rule VIII are left to the student. They are most readily arranged by expressing the given function in terms of sine and cosine and proceeding as above.

Proof of Rule IX (a). If $y = \sin^{-1} x$, then $\sin y = x$, and applying Rules V, VII, and VIII (a) we have,

$$\cos y \frac{dy}{dx} = 1, \text{ or } \frac{dy}{dx} = \frac{1}{\cos y} = \frac{\pm 1}{\sqrt{1 - \sin^2 y}} = \frac{\pm 1}{\sqrt{1 - x^2}}$$

Understanding the radical to have a positive value, we must prefix a negative sign in those quadrants in which the cosine is negative, the second and third quadrants.

The remaining parts of Rule IX are proved in similar manner, and the work is left to the student.

The rules given in this article, combined with Rule V, serve to find derivatives of all simple trigonometric and inverse trigonometric functions, and combined with the other rules that precede them they serve to find derivatives of functions expressed by algebra and trigonometry combined.

We illustrate by two examples.

Let $y = \tan (3x^2 - 6)$, then putting $3x^2 - 6 = u$, we have $y = \tan u$, and applying Rules V and VIII (c), we have

$$\frac{dy}{dx} = \frac{du}{dx} \sec^2 u = 6x \sec^2 (3x^2 - 6).$$

Let $y = x^2 \sin^{-1} \frac{x}{1 + x}$, then putting $\frac{x}{1 + x} = u$, $y = x^2 \sin^{-1} u$, and applying Rules II, V and IX (a) we obtain.

$$\frac{dy}{dx} = \sin^{-1} u \frac{d}{dx} (x^2) + x^2 \cdot \frac{1}{\sqrt{1 - u^2}} \frac{du}{dx}$$

and this becomes

$$2x \sin^{-1} \frac{x}{1 + x} + \frac{x^2}{\sqrt{1 - \frac{x^2}{(1 + x)^2}}} \cdot \frac{1}{(1 + x)^2}$$

or

$$\frac{dy}{dx} = 2x \sin^{-1} \frac{x}{1+x} + \frac{x^2}{(1+x)\sqrt{1+2x}}$$

EXERCISES VIII

1. Find the derivative with respect to x of each of the following:

(a) $\cos 3x$; (f) $2 \sin x - 6 \cos 3x$;

(b) $\tan \frac{x}{2}$; (g) $a \sin 2x \cos 3x$;

(c) $\sin\left(x - \frac{\pi}{6}\right)$; (h) $b \tan x / \sin 2x$;

(d) $\cot\left(-\frac{x}{2}\right)$; (i) $\sin mx \cos nx$;

(e) $\sec\left(\frac{\pi}{3} - 2x\right)$; (j) $\cot mx \tan nx$.

2. Show that the derived function is the same for $-2 \sin^2 x$ as for $\cos 2x$. What conclusion can you draw from this fact as to the relation between the two given functions?

3. From the formula

$$\sin(mx + nx) = \sin mx \cos nx + \cos mx \sin nx$$

derive by differentiation the formula for $\cos(mx + nx)$.

4. Find the following derived functions:

(a) $\frac{d}{dt}(\sin^2 t)$; (d) $\frac{d}{dt}(t \sin 3t + b \sin^2 3t)$;

(b) $\frac{d}{d\theta}\cos^3\left(\frac{\pi}{4} - \theta\right)$; (e) $\frac{d}{dx}(1 - x^2 \sin 2x + \cos 3x)$;

(c) $\frac{d}{dx}(\cos^2 x \sin 2x)$; (f) $\frac{d}{d\theta}(\sec^2 \theta \tan \theta)$.

5. A wheel is rotating in a vertical plane at one revolution per second, i.e., at 2π radians per second. If the diameter of the wheel is 4 feet, with what speed is a point on the rim rising when it is on a level with the center? When it is a foot above the center?

6. Differentiate the following functions with respect to the variable present:

(a) $\sin^{-1} 2u$; (d) $x \tan^{-1} 2x$; (g) $v(\cot^{-1} v)^2$;

(b) $\cot^{-1} \frac{x}{2}$; (e) $\tan^{-1}(\cos x)$; (h) $\sin^{-1}(x^2 - 5)$;

(c) $\sec^{-1} \frac{1}{y}$; (f) $\tan(\cos^{-1} x)$; (i) $\frac{\sin^{-1} 3\theta}{\cos^{-1} 2\theta}$.

Art. 12. Derivatives of Exponential and Logarithmic Functions

The special rules required here are:

Rule X. (a) $\frac{d}{dx}(e^x) = e^x;$

(b) $\frac{d}{dx}(a^x) = a^x \log_e a.$

Rule XI. (a) $\frac{d}{dx}(\log_e x) = \frac{1}{x};$

(b) $\frac{d}{dx}(\log_a x) = \frac{1}{x}\log_a e.$

It will be observed that Rule X (a) is a special case of X (b), for if $a = e$, we have $\log_e e = 1$, so that (b) reduces to (a). Rule X (b) holds for any base a, while Rule X (a) holds only for the particular base which we denote by e. In Chapter I we were told the value of this number, the natural or naperian base, and its importance was suggested, but the facts which lead to the selection of this particular number as a base remain to be presented in the proof of the rule.

Proof of Rule X (a). First let us consider an exponential function to an arbitrary base, a:

$$y = a^x$$

Giving x an increment Δx, we have

$$y + \Delta y = a^{x+\Delta x}$$

therefore, subtracting, and dividing by Δx,

$$\frac{\Delta y}{\Delta x} = \frac{a^{x+\Delta x} - a^x}{\Delta x} = a^x\left(\frac{a^{\Delta x} - 1}{\Delta x}\right).$$

Taking the limit as Δx approaches zero, we have

$$\frac{dy}{dx} = a^x \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

It remains to discover the value of the last factor,

$$\lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

Now this unknown constant is clearly the value of the derivative at $x = 0$, since $a^0 = 1$; and to investigate its value we make graphs of $y = a^x$ for several values of a. (Fig. 16.) The derivative at $x = 0$ shows the rate of increase of y compared to that of x, that is, it shows the direction of the graph as it crosses the y-axis. This differs for different values of a, but it is obvious that for some value of a, between 2 and 3, the graph will cross at 45°, so that y is increasing at the same rate as x, and so, for that value of a, the derivative at $x = 0$ is 1. Let us denote that value of a by e, then we may write

$$\lim_{\Delta x \to 0} \frac{e^{\Delta x} - 1}{\Delta x} = 1$$

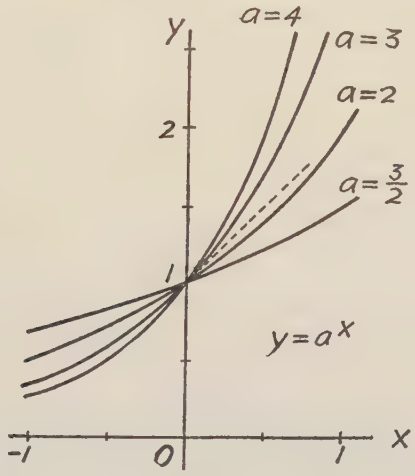


Fig. 16

and, therefore, if $y = e^x$, we have

$$\frac{dy}{dx} = e^x.$$

It will be shown later that the definition of e given here, viz., **that value of a for which the graph of $y = a^x$ crosses the y-axis at 45°**, yields a method for computing its value to any desired accuracy. For the present we will accept the value as stated in Art. 3, $e = 2.71828$.

Proof of Rule X (b). Starting with the function $y = a^x$, we proved above that for a certain special value, $a = e$, the derived function is equal to the original function. Now let $a = e^m$, then $m = \log_e a$, and

$$y = a^x = e^{mx}.$$

Applying Rule X (a) and Rule V, we have

$$\frac{dy}{dx} = e^{mx}\, m = a^x \log_e a.$$

Proof of Rules XI (a) and XI (b). Here again, (a) is a special case of (b), and follows from it by putting $a = e$. It is sufficient, therefore, to prove (b). Let

$$y = \log_a x,$$

then

$$a^y = x.$$

Taking the derivative of both sides with respect to x, using Rules X (b), V and VII, we have

$$a^y \log_e a \frac{dy}{dx} = 1.$$

Solving for $\frac{dy}{dx}$, remembering that $a^y = x$ and $\log_e a = \frac{1}{\log_a e}$ (since if $e^m = a$, $a^{\frac{1}{m}} = e$), we obtain

$$\frac{dy}{dx} = \frac{1}{a^y \log_e a} = \frac{1}{x} \log_a e.$$

Besides the natural base, e, the base most used is the **common** base 10. For numerical computations the common base possesses important advantages, but for manipulation of functions in the various mathematical processes the natural base is preferable, and any other base, a, must be referred to it by means of the **modulus,** $\log_a e$. The modulus of the common base, $\log_{10} e$, may be taken from a table of common logarithms, its value to five places being 0.43429. Unless otherwise indicated all logarithmic functions will be understood as to the base e.

Rules I to XI are sufficient to find the derivative of any function that may be expressed in terms of the elementary functions that we have considered. Facility in the use of these rules comes only with persistent practice. It is considered profitable to memorize most, if not all, of the rules given, as well as many formulas that result from these rules when applied to functions of special types. In this the students further experience must be his guide, but the exercises of the text are selected to afford familiarity with the types that are most important.

EXERCISES IX

1. Find derivatives of the following functions:

(a) $4e^{2x}$; (e) $\log(3x - 2)$; (i) $e^{\sin x}$;

(b) $e^{-\frac{1}{2}x}$; (f) $2\log(x^2 + 4)$; (j) $\log\cos x$;

(c) a^{x^2-7}; (g) $\log_{10}(4x)$; (k) $a^{\log_a x}$;

(d) $3a^{-\frac{2}{x}}$; (h) $7\log_{10}(1 - 2x)$; (l) $\log_a(a^x)$.

2. Obtain the following derived functions:

(a) $\frac{d}{dx}(xe^x - e^{-x})$; (d) $\frac{d}{dx}(x^2 - 2x)\log(1 - x)^2$;

(b) $\frac{d}{du}e^u \sin u$; (e) $\frac{d}{dt}\frac{\log(1 - t^2)}{2t - 3}$;

(c) $\frac{d}{d\theta}a^\theta \tan\left(\frac{\theta}{2} - \frac{\pi}{4}\right)$; (f) $\frac{d}{dt}\sin\left(2r - \frac{\pi}{6}\right)\log\frac{r}{1 - r}$.

3. Find the derivative of the Gudermannian of x,

$$\text{gd}x = \tan^{-1}\frac{e^x - e^{-x}}{2}.$$

4. A class of functions, called hyperbolic functions, are defined in terms of exponential functions:

Hyperbolic sine of x: $\sinh x = \frac{e^x - e^{-x}}{2}$

Hyperbolic cosine of x: $\cosh x = \frac{e^x + e^{-x}}{2}$

Hyperbolic tangent of x: $\tanh x = \frac{\sinh x}{\cosh x}$

Hyperbolic cotangent of x: $\coth x = \frac{1}{\tanh x}$

Hyperbolic secant of x: $\text{sech } x = \frac{1}{\cosh x}$

Hyperbolic cosecant of x: $\text{csch } x = \frac{1}{\sinh x}$

Show that $\frac{d}{dx}\sinh x = \cosh x$ and $\frac{d}{dx}\cosh x = \sinh x$, and derive rules for differentiation of the other hyperbolic functions, observing similarity to rules for trigonometric functions.

5. Express in terms of exponentials the last four functions defined

in exercise 4, and by finding the derivatives of these expressions prove again the rules derived in exercise 4.

6. Understanding that $y = \sinh^{-1} x$ means the same as $x = \sinh y$, show, using the method by which Rule IX (a) was proved, that

$$\frac{d}{dx}\sinh^{-1}x = \frac{d}{dx}\log(x + \sqrt{x^2 + 1}) = \frac{1}{\sqrt{x^2 + 1}}.$$

Similarly show that

$$\frac{d}{dx}\tanh^{-1}x = \frac{d}{dx}\frac{1}{2}\log\frac{1 + x}{1 - x} = \frac{1}{1 - x^2}.$$

7. Show that the derivative of any function of the form

$$y = e^{kx}(a \sin nx + b \cos nx)$$

is another function of the same form, but with different values for a and b, in general.

8. Show that the derivative of any function of the form

$$y = e^{kx}(a_0x^n + a_1x^{n-1} + \ldots + a_n)$$

is another function of the same form, but with different values, in general, for the coefficients of the polynomial.

Art. 13. Functions Defined Indirectly

We have already considered the case where y is a function of an an auxiliary variable u, while u is a function of x, and we have seen how to obtain the graph of y as a function of x, as well as how to find the derivative of y with respect to x. This is, of course, an indirect method of expressing y as a function of x, but since the elimination of u involves a mere substitution of its value in terms of x, functions y so defined have been treated along with functions directly expressed in terms of the independent variable, called explicit functions. We now give attention to two indirect modes of expressing a functional relation between two variables.

(1) **Implicit functions.** This is the name given to functions defined by an equation which involves both independent and dependent variables, but which is not solved for the dependent variable. If the equation can be solved for the dependent variable, we may thus obtain the function in explicit form, but it is often more convenient to treat the function in the implicit form. Taking x and y as independent

and dependent variables, respectively, we may suppose every term of the defining equation transposed to the left side, so that we have an equation of the form

$$f(x, y) = 0,$$

i.e., an expression involving both variables equated to zero. To find the derivative of y with respect to x we have only to differentiate the equation with respect to x, remembering that y is a function of x, then solve for the derivative, obtaining an equation of the form

$$\frac{dy}{dx} = \phi(x, y),$$

expressing the derivative as a function of both variables. In using this, we must remember that the two variables are connected by the relation expressed in the original equation.

For example, let

$$x^2y + x \sin^{-1}\left(\frac{y}{x}\right) = 0.$$

This equation cannot be solved for y, but, taking the derivative, treating each term as a product, we have

$$x^2\frac{dy}{dx} + 2xy + x\frac{d}{dx}\sin^{-1}\left(\frac{y}{x}\right) + \sin^{-1}\left(\frac{y}{x}\right) = 0,$$

or

$$x^2\frac{dy}{dx} + 2xy + \frac{x\dfrac{dy}{dx} - y}{\sqrt{x^2-y^2}} + \sin^{-1}\left(\frac{y}{x}\right) = 0,$$

from which

$$\frac{dy}{dx} = -\frac{\left[2xy + \sin^{-1}\left(\dfrac{y}{x}\right)\right]\sqrt{x^2 - y^2} - y}{x^2\sqrt{x^2 - y^2} + x}.$$

Taking any values of x and y which satisfy the original equation, we may substitute their values in this last expression and obtain the corresponding value of the derivative.

(2) **Parametric equations.** Here an auxiliary variable is used, but instead of having this variable expressed as a function of the inde-

pendent variable, we have both independent and dependent variables expressed in terms of the auxiliary variable.

Now Rule V states that if $y = f(u)$ and $u = \phi(x)$, then

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx'}$$

and if $\frac{du}{dx}$ is not zero, we may divide by it, obtaining

$$\frac{\frac{dy}{dx}}{\frac{du}{dx}} = \frac{dy}{du}.$$

Here we may legitimately regard both y and u as functions of x.

For the parametric equations let us take t as the auxiliary variable, or parameter, and write

$$y = f(t), \quad x = \phi(t),$$

then, using the principle just developed,

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}}.$$

Often the auxiliary variable used as parameter has vital physical or geometrical significance, so that equations in parametric form, with the derivative expressed in terms of the parameter, afford information that could not be obtained from the (x, y) equation found by eliminating the parameter.

As a simple example, let us find the derivative of the volume of a sphere with respect to the surface. We have

$$v = \frac{4}{3}\pi r^3, \quad s = 4\pi r^2,$$

therefore,

$$\frac{dv}{ds} = \frac{\frac{dv}{dr}}{\frac{ds}{dr}} = \frac{4\pi r^2}{8\pi r} = \frac{1}{2}r.$$

If we had expressed v directly as a function of s we should have had radicals to treat, and the result would not have been as easily interpreted. In the present form we see that in a small sphere the surface changes much faster than the volume, while in a large sphere the volume changes faster than the surface.

EXERCISES X

1. From each of the following equations find the derivative of y with respect to x as a function of x and y:

(a) $x^2 + y^2 = 16$;

(b) $xy - x^2 + 2xy^2 = 4$;

(c) $(x^2 - 6)(y^3 - x) = 2xy$;

(d) $(x - y)(x + y - 2) = 1$;

(e) $\tan^{-1}\frac{y}{x} = x$;

(f) $\cos(x - y) = 0$;

(g) $x \sinh y = y$;

(h) $e^x \log y = 1$.

2. In the following cases find, in terms of the parameter, the derivative of the first variable with respect to the second:

(a) $y = \frac{1}{1 - t}$, $x = \frac{1}{1 + t}$;

(b) $u = 1 - \sin\theta$, $v = \tan\theta$;

(c) $y = 1 - t^3$, $x = t^2 + 1$;

(d) $y = \frac{1}{2}\sqrt{2t^3}$, $x = \frac{1}{3}t^2$;

(e) $y = a\sin\theta + b\cos\theta$, $x = a\cos\theta + b\sin\theta$;

(f) $v = \frac{2}{3}\pi r^3$, $s = \pi(1 + \sqrt{2})\, r^2$.

3. If a wheel rolls on a horizontal track, θ being the angle turned through, then the horizontal distance x, and the vertical distance y, from its starting point of a point on the wheel that was initially in contact with the tract, are given by

$$x = a(\theta - \sin\theta), \quad y = a(1 - \cos\theta),$$

where a is the radius of the wheel. What is the direction of motion of the point when $\theta = \frac{\pi}{4}$? When $\theta = \frac{\pi}{2}$? When $\theta = \pi$? If the wheel turns at a steady rate of one quarter turn per second, how fast are x and y increasing when $\theta = \frac{\pi}{2}$? When $\theta = \frac{3\pi}{2}$?

Art. 14. Successive Derivatives

We have seen that the derivative of a function with respect to the independent variable, when derived without specifying the value of the variable, is a new function of the variable, called the derived function. We have learned the processes for finding the derivatives of all elementary types of functions. It requires no new theory to take the derivative of a derived function, obtaining the **second derivative,** or **second derived function.** This function may again be differentiated, obtaining the **third derivative,** of the original function, and so on indefinitely. Many practical problems require the use of successive derivatives, and still others demand the conception of certain known functions as one of the higher derivatives of an unknown function that is to be determined. Thus the acceleration of a moving object is the derivative of the velocity with respect to time, or the second derivative of the displacement with respect to time. The deflection at any point of a loaded beam may be regarded as a function of the distance from one end. The second derivative of this function with respect to the distance is proportional to the bending moment, the third derivative to the shearing force, and the fourth derivative to the load intensity at the given position.

Given a function $y = f(x)$, we have used the notation

$$\frac{dy}{dx} = \frac{d}{dx}f(x)$$

to denote the derived function, and this would give for the second derivative,

$$\frac{d}{dx}\frac{dy}{dx} = \frac{d}{dx}\frac{d}{dx}f(x).$$

A convenient contraction of this notation is

$$\frac{d^2y}{dx^2} = \frac{d^2}{dx^2}f(x),$$

indicating that the operation $\frac{d}{dx}$, of differentiating with respect to x, is applied twice, and not, of course, that anything has been squared. In this way we have for the successive derivations of y

$$\frac{dy}{dx}, \frac{d^2y}{dx^2}, \frac{d^3y}{dx^3}, \cdots\cdots \frac{d^ny}{dx^n},$$

and if $y = f(x)$ we may use, with the same meaning,

$$\frac{d}{dx}f(x),\ \frac{d^2}{dx^2}f(x),\ \ldots\ldots\ \frac{d^n}{dx^n}f(x).$$

Other notations sometimes used for the same purpose are

$$D_x y,\ D_x^2 y,\ D_x^3 y,\ \ldots\ldots D_x^n y;$$

$$y',\ y'',\ y',\ y''',\ \ldots\ldots y^{(n)};$$

$$f'(x),\ f''(x),\ f'''(x),\ \ldots\ldots f^{(n)}(x).$$

If y is given explicitly in terms of x, the successive derivatives may be obtained directly, but if y is indirectly defined as a function of x the process of finding higher derivatives may be somewhat involved with elimination of derivatives of lower order that recur in the successive differentiations. To illustrate, let it be required to find the first and second derivatives of y with respect to x from the equation

$$x^2y + y^3 = 6.$$

Differentiating once we have

$$x^2\frac{dy}{dx} + 2xy + 3y^2\frac{dy}{dx} = 0,$$

or

$$\frac{dy}{dx} = -\frac{2xy}{x^2 + 3y^2}.$$

Differentiating again,

$$\frac{d^2y}{dx^2} = -\frac{(x^2 + 3y^2)\left(2x\frac{dy}{dx} + 2y\right) - 2xy\left(2x + 6y\frac{dy}{dx}\right)}{(x^2 + 3y^2)^2}$$

and substituting in this the value of $\frac{dy}{dx}$, and reducing,

$$\frac{d^2y}{dx^2} = -\frac{2y(x^2 - 3y^2)}{(x^2 + 3y^2)^3}.$$

It must be remembered that this expression for the second derivative requires that y be related to x by the original equation.

An alternative method is to write the original equation and differentiate twice before solving, thus:

$$x^2y + y^3 = 6,$$

$$(x^2 + 3y^2)\frac{dy}{dx} + 2xy = 0,$$

$$(x^2 + 3y^2)\frac{d^2y}{dx^2} + \left(2x + 6y\frac{dy}{dx}\right)\frac{dy}{dx} + 2x\frac{dy}{dx} + 2y = 0.$$

These three equations are sufficient to determine y, $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ as functions of x, or we may use the last two to determine the two derivatives as functions of x and y, regarding y as determined by the first equation.

If the function is defined by means of parametric equations the problem of elimination arises in slightly different manner. Let

$$y = t^2, \quad x = \frac{1}{3}t^3,$$

then

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{2t}{t^2} = \frac{2}{t}.$$

To find the second derivative we have, since $\frac{dt}{dx} = \frac{1}{\frac{dx}{dt}} = \frac{1}{t^2}$,

$$\frac{d^2y}{dx^2} = \frac{d}{dt}\left(\frac{dy}{dx}\right)\frac{dt}{dx} = -\frac{2}{t^2}\cdot\frac{1}{t^2} = -\frac{2}{t^4}.$$

An alternative method is to differentiate both equations twice with respect to x:

$$y = t^2, \qquad x = \frac{1}{3}t^3,$$

$$\frac{dy}{dx} = 2t\frac{dt}{dx}, \qquad 1 = t^2\frac{dt}{dx},$$

$$\frac{d^2y}{dx^2} = 2\left(\frac{dt}{dx}\right)^2 + 2t\frac{d^2t}{dx^2}; \quad 0 = 2t\left(\frac{dt}{dx}\right)^2 + t^2\frac{d^2t}{dx^2}.$$

From these equations we may solve algebraically for $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ as functions of t, eliminating $\frac{dt}{dx}$ and $\frac{d^2t}{dx^2}$.

EXERCISES XI

1. Find the first, second and third derivatives of the following functions:

(a) $\sin 3x$; (c) e^{-2x}; (e) $\frac{1}{1-x}$;

(b) $\log (1 - x^2)$; (d) $\tan \frac{x}{2}$; (f) $x \cos 2x$.

2. Find $\frac{d^2\theta}{dr^2}$ from the equation

$$r^2\theta = \sin \theta.$$

3. If $x = a \cos \theta$ and $y = a \sin \theta$, show that

$$\left[1 + \left(\frac{dy}{dx}\right)^2\right]^3 = a^2 \left(\frac{d^2y}{dx^2}\right)^2.$$

4. Show that the successive derivatives of a function of the form

$$y = (a + bx) \sin x + (c + dx) \cos x$$

are all functions of the same form, differing only in the values of the constant coefficients.

5. Show that the relation

$$\frac{d^3y}{dx^3} + 3\frac{d^2y}{dx^2} + \frac{dy}{dx} = -3y$$

is fulfilled if y is any one of the following functions: $\sin x$; $\cos x$; e^{-3x}.

6. If a body moves so that its distance s in feet is given in terms of the time t in seconds by any of the following equations, show that the acceleration is 20 feet per second per second.

$$s = 10t^2; \quad s = 10t^2 - 17t; \quad s = 10t^2 + 3t + 80.$$

7. Find $\frac{d^2y}{dx^2}$ from the equations

$$x = a\theta - b \sin \theta, \; y = a - b \cos \theta.$$

CHAPTER IV

INTEGRALS OF FUNCTIONS

Art. 15. Inverse of Differentiation

In Chapter III we have considered the process of finding the derivative of a function, and we understand that this derivative is a new function, the derived function. Here we give attention to the inverse problem, that of finding a function when its derivative, or derived function, is given.

This process gives rise to a new notation. Using the notation of the last chapter we may write, if $y = f(x)$,

$$\frac{dy}{dx} = \frac{d}{dx} f(x) = f'(x).$$

This indicates that an operation has been applied to y, or $f(x)$, to obtain the derived function, $f'(x)$. Now to denote the inverse process we write

$$\int f'(x)dx = f(x),$$

the symbol $\int \ldots dx$ indicating that an operation has been applied to $f'(x)$ to obtain $f(x)$. This operation, or process, the inverse of differentiation, is called **integration,** and the resulting function, $f(x)$, is called the **integral, with respect to** x, **of** $f'(x)$. Here it is natural to think of $f'(x)$ as the original function, known as the **integrand,** and it is not usual to use the accent, which denotes its character as derivative. Thus, given the functions

$$f(x),\ \phi(x),\ F(x), \text{ etc.},$$

we indicate their integrals by

$$\int f(x)dx,\ \int \phi(x)dx,\ \int F(x)dx, \text{ etc.},$$

respectively.

Since the derivative of any constant is zero, any two functions that differ only by a constant term have the same derivative. Therefore, when we find an integral of any function, any constant may be added to the integral and it still remains an integral of the given function. Thus, the **general integral** of any function contains an arbitrary constant as one term, called the **constant of integration.**

For example, since the functions $x^2 - 7$, $x^2 + 2$, $x^2 + \sqrt{3}$, $x^2 + \frac{\pi}{6}$, etc., all have the derivative $2x$, we have

$$\int 2x\,dx = x^2 + c,$$

where c, the constant of integration, may have any constant value whatever.

We have seen that finding the derivative of a function is equivalent to finding the rate of change of the value of the function in terms of the rate of change of the independent variable. Now the problem of integration arises when we know a function which expresses the rate of increase of another function that we wish to find. To illustrate, let us assume that the acceleration, or increase of velocity, of any freely falling body is 32 feet per second per second downward. Then the velocity v in feet per second is a function of the time in seconds, given by the integral

$$v = \int 32\,dt = 32t + v_o.$$

Here v_o is the constant of integration, and this particular notation is given to it because it is obviously the value of v when $t = 0$. Now velocity is the rate of increase of distance, therefore the distance s, in feet, is given by the integral

$$s = \int (32t + v_o)dt = 16t^2 + v_o t + s_o.$$

The new constant of integration is given the notation s_o because it is the value of s when $t = 0$. If the initial values v_o and s_o of v and s are known and substituted, then the integrals are not general, but they are known as **particular integrals** corresponding to the **given initial conditions.**

In Article 8, considering the graph of a function, $y = f(x)$, we saw that the value of the derivative at any value of x is determined by the

direction of the tangent line to the graph at that value of x. Now when we are given the derived function, f'(x), we are given the direction of the graph of the integral function at every value of x, and it is clear that this determines the graph, called the **integral curve,** except for the arbitrary constant. A change of the constant merely shifts the curve upward or downward, changing all ordinates by a certain quantity, but not changing the direction of the graph at any value of x, as at x_1 (Fig. 17).

Additional knowledge of the significance of integration is obtained by consideration of the graph, $y = f'(x)$ (Fig. 18).

Let us take a function $A = f(x)$, such that the value of A at any value of x, as x_1, is the area between the curve $y = f'(x)$ and the x-

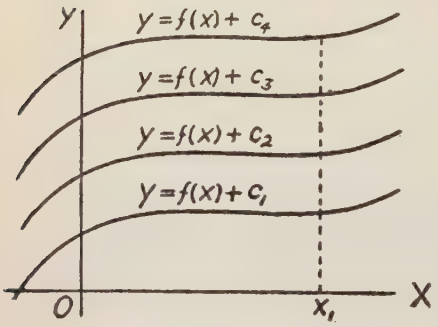


Fig. 17

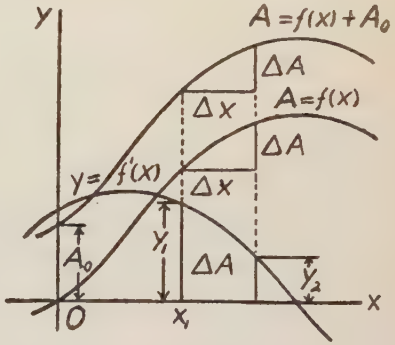


Fig. 18

axis and between the ordinates at $x = 0$ and $x = x_1$. We will now show that the derivative of A with respect to x is f'(x), and therefore that

$$A = f(x) = \int y\,dx = \int f'(x)\,dx.$$

Consider two values of x, x_1 and $x_1 + \Delta x$, and the corresponding ordinates y_1 and y_2. The area, ΔA, between the ordinates at these two places is the increment in A corresponding to the increment Δx in x. Now the derivative of A with respect to x is

$$\frac{dA}{dx} = \lim_{x \to 0} \frac{\Delta A}{\Delta x}.$$

But $\Delta A/\Delta x$ is less than the greatest ordinate between y_1 and y_2, and greater than the least ordinate. As Δx approaches zero, y_2 ap-

proaches y_1, assuming that the graph is continuous, and the greatest and least ordinate approach coincidence at y_1; therefore the ratio $\Delta A/\Delta x$ approaches y_1. The conclusion is that, at any x, we have

$$\frac{dA}{dx} = \frac{d}{dx} f(x) = y = f'(x).$$

The rate at which the graph of $A = f(x)$ rises as x increases is measured by the ordinate of the graph of $y = f'(x)$, and when the ordinate of this curve is negative we find the value of A decreasing. The function $A = f(x) + A_o$, where A_o is any constant, has also the derivative $f'(x)$ and is therefore an integral of this function.

In the proof given above it is assumed that the graph of $y = f'(x)$ is continuous. For the present we must permit this restriction, but it will be shown later that certain kinds of discontinuous functions have integral functions that are continuous. This means of course, that a continuous function exists whose derivative is the given discontinuous function.

Familiarity with the process of finding derivatives enables us, without special formulas, to perform certain simple integrations. The following exercises should serve to fix the ideas as to the meaning of integration.

EXERCISES XII

1. Write the integrals of the following functions, remembering in each case to add the constant of integration. Test your result by comparing its derivative with the given function:

(a) $x^3 + 3$; (c) $\sin x$; (e) e^{3x};

(b) $2x - x^2$; (d) $\cos 2x$; (f) $4e^{-2x}$.

2. Find the general integral of the function

$$y = 3x^2 - 4x + 2,$$

and find the particular integral which has the value of 10 when $x = 1$.

3. Obtain the expression for the following integral which pertains to the initial condition $y = 6$ when $x = 0$,

$$y = \int 5 \cos x \, dx.$$

4. If an object is moving in such manner that its speed v in feet per second is expressed in terms of the time t in seconds by $v =$

$12 - 6t$, find how far it will move during the time from $t = 0$ to $t = 2$; from $t = 2$ to $t = 5$; from $t = 0$ to $t = 5$.

5. Suppose it known that $u = 2$ when $v = 4$, and that the rate of change of u in terms of the rate of change of v is a function of v, given by the equation

$$\frac{du}{dv} = 1 - 6v^2.$$

Find u as a function of v.

6. Find the area between the graph of $y = \sin x$ and the x-axis, from $x = 0$ to $x = \pi$.

7. Find as a function of x, the area between the graph of $y = e^{-x}$ and the x-axis and between the y-axis and the ordinate at x.

8. Water is flowing into a tank at the rate of one barrel per minute. Express the quantity of water in the tank at time t as an integral.

9. If $\frac{d^2y}{dx^2} = 6x - 2$, what is the general expression for y as a function of x? If $y = 2$ and $\frac{dy}{dx} = 0$ when $x = 0$, what are the values of y and $\frac{dy}{dx}$ at $x = 2$? At $x = 5$? At $x = -2$?

Art. 16. Formulas for Integration

In Chapter III rules for differentiation were developed that are sufficient to differentiate any elementary function for which a derivative exists. Now integration is an indirect process, and while it may be shown that all continuous functions and some that are not continuous possess integrals, it frequently happens that these can not be expressed in terms of elementary functions, and when they can the only means of finding them consists essentially in reducing the integrand to a form which is recognizable as the derivative of some known function. The need for recognizing these relations gives rise to a group of **standard forms** of integrals, each based on one of the well known rules for differentiation. In these forms u and v are used as before to denote functions of x. From Rule I we have,

$$1. \int (u + v)dx = \int udx + \int vdx.$$

The full application of Rule II to the problem of integration is deferred

to Chapter VI, but from the special case when one of the factors is constant we have,

2. $\int a\ u\ dx = a \int u dx.$

By considering the case when $u = x$ in Rule III, we obtain the following, in case $n \neq -1$ (see 17 below):

3. $\int x^n dx = \frac{x^{n+1}}{n+1} + c.$

From Rules VI and VII and 2 above, we have,

4. If $\frac{dy}{dx} = 0$, then $y =$ constant,

5. $\int a\ dx = ax + c.$

The following forms, based primarily on the several parts of Rules VIII, IX and XI may be verified directly by differentiation:

6. $\int \cos x\ dx = \sin x + c.$

7. $\int \sin x\ dx = -\cos x + c.$

8. $\int \sec^2 x\ dx = \tan x + c.$

9. $\int \csc^2 x\ dx = -\cot x + c.$

10. $\int \sec x \tan x\ dx = \sec x + c.$

11. $\int \csc x \cot x\ dx = -\csc x + c.$

12. $\int \frac{1}{\sqrt{1-x^2}} dx = \pm \sin^{-1} x + c_1 = \mp \cos^{-1} x + c_2.$

13. $\int \frac{1}{1+x^2} dx = \tan^{-1} x + c_1 = -\cot^{-1} x + c_2.$

14. $\int \frac{1}{x\sqrt{x^2-1}} dx = \pm \sec^{-1} x + c_1 = \mp \csc^{-1} x + c_2.$

15. $\int e^x dx = e^x + c.$

16. $\int a^x dx = a^x \log_a e + c.$

17. $\int \frac{1}{x} dx = \log x + c.$

From the relations given in Exercise 6, of IX, and similar relations involving $\cosh^{-1}x$, we have three additional forms:

18. $\int \frac{1}{\sqrt{x^2+1}} dx = \log(x + \sqrt{x^2+1}) + c = \sinh^{-1}x + c.$

19. $\int \frac{1}{\sqrt{x^2-1}} dx = \log(x + \sqrt{x^2-1}) + c = \cosh^{-1}x + c.$

20. $\int \frac{1}{1-x^2} dx = \frac{1}{2} \log \frac{1+x}{1-x} + c = \tanh^{-1}x + c.$

The usefulness of these forms, 1–20, is greatly increased by application of a principle derived from Rule V. In the statement of Rule V it is understood that y is a function of u while u is a function of x, so that y is a function x that may be obtained by replacing u by its value in terms of x. Let us assign the notations,

$$y = f(u),\ u = \phi(x),\ y = f[\phi(x)] = F(x).$$

Now by the rule we have,

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} \text{ or } F'(x) = f'(u) \cdot \phi'(x).$$

We see that

$$\int F'(x)dx = \int f'(u) \cdot \phi'(x)dx = F(x) = y.$$

Also we see, treating u as independent variable,

$$\int f'(u)du = f(u) = y = F(x).$$

The great advantage we obtain from this lies in the fact that if we have an integrand, as $F'(x)$, which we are unable to recognize as a derivative of any function, we may separate it into two factors, one of which is a function of a quantity u, as $f'(u)$, while the other factor is the derivative of u. Then we have only to consider the integral

$$\int f'(u)du,$$

which, factors being wisely chosen, may be easily recognizable. The device used here is known as the **method of substitution.**

To illustrate, consider the integral

$$x\sqrt{1+x^2}\,dx$$

Put $u = 1 + x^2$, then $\frac{du}{dx} = 2x$, so that,

$$\int x\sqrt{1+x^2}\,dx = \frac{1}{2}\int (1+x^2)^{\frac{1}{2}} \cdot 2x\,dx = \frac{1}{2}\int u^{\frac{1}{2}}\frac{du}{dx}dx.$$

Now by the method of substitution we may integrate as follows, using forms 2 and 3:

$$\frac{1}{2}\int u^{\frac{1}{2}}du = \frac{1}{3}u^{\frac{3}{2}} + c = \frac{1}{3}(1+x^2)^{\frac{3}{2}} + c.$$

Again, let us consider the integral,

$$\int \tan x\,dx = \int \frac{\sin x}{\cos x}\,dx.$$

Put $u = \cos x$, then $\frac{du}{dx} = -\sin x$, and we may write, using form 17:

$$\int \tan x\,dx = -\int \frac{1}{u}\frac{du}{dx}\,dx = -\int \frac{1}{u}du = -\log u$$

$$= \log \frac{1}{u} = \log \frac{1}{\cos x} = \log \sec x.$$

EXERCISES XIII

1. Show that $\int \frac{1}{\sqrt{4-9x^2}}dx = \frac{1}{3}\sin^{-1}\frac{3x}{2} + c.$

Suggestion: Put $u = \frac{3}{2}x$ and use form 12.

2. Show that $\int \frac{3}{1+4x^2}dx = \frac{3}{2}\tan^{-1}2x + c.$

3. Show that $\int \sin x \cos x\,dx = \frac{1}{2}\sin^2 x + c.$

4. Find the following integrals:

(a) $\int \cos 4x\, dx$; (c) $\int \sin\left(\frac{\pi}{2} - x\right) dx$;

(b) $\int e^{3x} dx$; (d) $\int e^{-\frac{x}{2}} dx$.

5. Find integral functions of the following functions:

(a) $\cot x$; (b) $\frac{x^2}{1+x^3}$; (c) $\sec^2 x \tan x$; (d) $x(1-x^2)^3$.

6. Given

$$\frac{dy}{dx} = \frac{x^3 + 3x^2 - 6x + 2}{1 - x^2}.$$

find y as a function of x.

Suggestion: Write, after finding quotient and remainder by division,

$$\frac{dy}{dx} = -(x+3) - \frac{5(1-x)}{x^2-1} = -x - 3 + \frac{5}{x+1};$$

whence

$$y = -\int x\, dx - \int 3dx + 5\int \frac{1}{x+1} dx.$$

7. Show that

$$\int \frac{3x-5}{4x^2+9} dx = \frac{3}{8}\log(4x^2+9) - \frac{5}{6}\tan^{-1}\left(\frac{2}{3}x\right) + c.$$

8. Find $\int \frac{1}{\sqrt{x^2 - a^2}} dx$.

9. Find $\int \cos^n x \sin x\, dx$.

10. Find $\int x^2 e^{x^3} dx$.

11. Find $\int \frac{1}{x\sqrt{x^2-4}} dx$.

12. Find $\int \frac{1}{x^2 + 4x - 5} dx$.

Hint: Write $x^2 + 4x - 5 = (x+2)^2 - 9$, then put

$$u = \frac{x+2}{3}.$$

Art. 17. The Definite Integral

When we wish to know how much the value of a function f(x) changes as x changes from one value to another, as from a to b, we have only to take the difference of the values of the function at the two places,

$$f(b) - f(a).$$

It often happens, as in examples 4 and 6 of Exercises XII, that we wish to find the change in value of a function when the rate of change, or derivative of the function is known. Denoting this derivative by f'(x), we have

$$f(x) = \int f'(x)dx.$$

This last expression involves a constant of integration, so the function values at x = a and x = b are not known, but since the constant does not change, the change of value of the function between x = a and x = b does not depend on the constant. For example, if

$$f'(x) = \cos 2x,$$

we may find the change in f(x) between x = 0 and x = $\frac{\pi}{4}$ by finding the integral

$$\int \cos 2x\, dx = \frac{1}{2}\sin 2x + c,$$

and taking values at 0 and at $\frac{\pi}{4}$ and subtracting. We have

$$\left(\frac{1}{2}\sin\frac{\pi}{2} + c\right) - \left(\frac{1}{2}\sin(0) + c\right) = \frac{1}{2} + c - c = \frac{1}{2}.$$

The standard notation for this operation is

$$\int_0^{\frac{\pi}{4}} \cos 2x\, dx = \left[\frac{1}{2}\sin 2x\right]_0^{\frac{\pi}{4}} = \frac{1}{2}.$$

Since the constant c does not influence the result, it is omitted from the expression in brackets.

In general, if we are given the derivative f'(x) of a function and we

wish to find how much the function changes from x = a to x = b, we write,

$$\int_a^b f'(x)dx = \Big[f(x)\Big]_a^b = f(b) - f(a).$$

The expression is known as a **definite integral,** and is, of course, a number, not a function. The numbers a and b are called respectively the **lower** and **upper limits of integration.** The function f'(x) is said to be **integrated from a to b.**

It is obvious that

$$\int_a^b \phi(x)dx = -\int_b^a \phi(x)dx,$$

i.e., a change of order of the limits is equivalent to a change of sign of the definite integral.

Also we have

$$\int_a^b \phi(x)dx + \int_b^c \phi(x)dx = \int_a^c \phi(x)dx,$$

which follows from the identical relation

$$[f(b) - f(a)] + [f(c) - f(b)] = f(c) - f(a),$$

where $f'(x) = \phi(x)$.

In distinction from the term definite integral, the integral function discussed in Article 15 is known as the **indefinite integral.** As we saw, the general expression for the indefinite integral is a function of x involving an arbitrary constant as one term. When we know the value of the function at one value of x, the value of the constant may be found, thus obtaining a **particular integral.** The notation for definite integral may be used to indicate the value of the constant of integration, by specifying the lower limit and using the upper limit as a variable. Thus

$$\int_a^x f'(x)dx = \Big[f(x)\Big]_a^x = f(x) - f(a).$$

Here the constant of integration takes the particular form $-f(a)$ and the result is equivalent to specifying that the integral function is zero when x = a.

As an illustration, suppose that

$$\frac{dy}{dx} = \frac{6}{x}, \text{ and } y = 0 \text{ when } x = 5.$$

We have then,

$$y = \int \frac{6}{x}\,dx = 6 \log x + c.$$

Using the initial conditions given we have $0 = 6 \log 5 + c$, or $c = -6 \log 5$, therefore

$$y = 6 \log x - 6 \log 5 = 6 \log \frac{x}{5}.$$

Now we may employ the new notation to condense this process,

$$y = \int_5^x \frac{6}{x}\,dx = \Big[6 \log x\Big]_5^x = 6 \log x - 6 \log 5 = 6 \log \frac{x}{5}.$$

If the initial conditions had given y any value other than zero, as $y = 8$ when $x = 5$, we should have written

$$y = 8 + \int_5^x \frac{6}{x}\,dx = \ldots = 8 + 6 \log \frac{x}{5}.$$

EXERCISES XIV

1. If $\frac{dy}{dx} = 3x^2 - 8x + 2$, how much does y change while x changes from 2 to 5?

2. Evaluate the definite integrals:

(a) $\int_0^1 t\sqrt{t}\,dt$; (c) $\int_1^5 \frac{\log x}{x}\,dx$;

(b) $\int_{-2}^0 2(t^2 - 1)\,dt$; (d) $\int_1^{-8} s^{\frac{2}{3}}(s^2 - 2s)ds$.

3. If an object moves under the action of a force, the work u, done by the force F, is a function of the distance s that the object moves, and the derivative of this function with respect to s, is an expression for the force as a function of s. A certain object is acted upon by a force $F = -4s$; how much work is done while s changes from 2 to 4? How much from 2 to -4? From -4 to 4?

4. A certain object is attracted toward a certain point by a force equal to $\frac{10}{s^2}$, where s is the distance from the point. How much work is done on the object by this force while the object moves from $s = 3$ to $s = 2$? From $s = 2$ to $s = 1$?

5. The velocity v at time t of a certain object is given by the equation

$$v = 4t - \sin 2t.$$

Express as an integral, with variable upper limit of integration the distance the object has moved at any time from its position at $t = \frac{\pi}{4}$. Express the value of this integral as a function of t.

Art. 18. Limit of a Sum of Differentials

Let us consider a function, $y = f(x)$, and its derived function $y' = f'(x)$. Let the curves in Figure 19 represent the graphs of these functions, then as we saw in Article 15, the change in the value of y as x changes from one value to another, as from a to b, is the area between the graph $y' = f'(x)$ and the x-axis and between the ordinates at a and b. Now let us divide the interval from a to b into several parts, each of length Δx, erecting ordinates at each division point, showing the value of y' at that point. On each interval of length Δx construct a rectangle with altitude equal to the left hand ordinate on the interval, then the area of each rectangle is $y'\Delta x = f'(x)\ \Delta x$ where y' or $f'(x)$ is the length of the left hand ordinate. Looking at the graph of $y = f(x)$ we see that the vertical line, marked dy, bears a ratio to Δx determined by the tangent line at the left extreme of Δx, so that

$$\frac{dy}{\Delta x} = y' = f'(x), \text{ or } dy = f'(x)\ \Delta x.$$

The length of the line dy, then, is equal to the area of the rectangle marked dy. This quantity dy is called the **differential of y,** and its value depends on the value of Δx and also on the value of x, $f'(x)$ being a function of x.

The sum of all the differentials from a to b may be denoted by

$$\sum_a^b dy = \sum_a^b f'(x)\ \Delta x,$$

and this sum differs from the area under the graph of $y' = f'(x)$ between a and b by a quantity which we may denote by E, the sum of

the areas between the upper bases of the rectangles and the curve. Let $\Delta y'$ be the greatest difference between values of y' on the base of any rectangle, then it is evident that

$$E < (b - a)\,\Delta y'.$$

Now nothing has been said concerning the magnitude of Δx, nor concerning the number of divisions of the interval from a to b. All that is said applies however small Δx may become, or however many the divisions of the interval from a to b. As Δx becomes smaller, approaching zero as a limit, the number of divisions increasing indefinitely, the value of $\Delta y'$ must approach zero, providing $f'(x)$ is a continuous

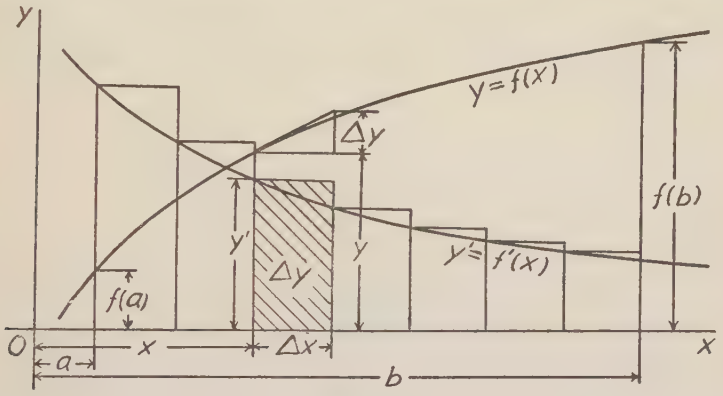


Fig. 19

function. This being true, E must approach zero, and the sum of the differentials approaches the area in question. We may write, then

$$\lim_{x \to 0} \sum_{a}^{b} f'(x)\,\Delta x = \int_a^b f'(x)\,dx = f(b) - f(a).$$

Considering the particular function $y = f(x) = x$, we have $y' = f'(x) = 1$, therefore $dy = dx = f'(x)\,\Delta x = \Delta x$, from which it is evident that for an independent variable there is no distinction between a differential and an increment. Accordingly, it is usual, when $y = f(x)$, to write

$$dy = \frac{dy}{dx}\,dx = f'(x)\,dx,$$

the differential, dx, of the independent variable being, in effect, an arbitrary increment, and the corresponding differential, dy, of the function being this increment multiplied by the derived function. The ratio of the differential of one variable to that of another is therefore the derivative of the first variable with respect to the other. It is from this point of view that the derivative is sometimes called the **differential coefficient.**

In discussing the limit of the sum of differentials above, we have, for the sake of simplicity, divided the interval from a to b into equal divisions, and have used the y′ at the left extreme of each division. With a few modifications of the argument which it would not be profitable to state here, it may be shown that even if the divisions of the interval are not equal, and if any y′ on each division is used, the limit of the sum as the width of the greatest Δx approaches zero is still equal to the definite integral. In fact, a glance at the figure will show that the

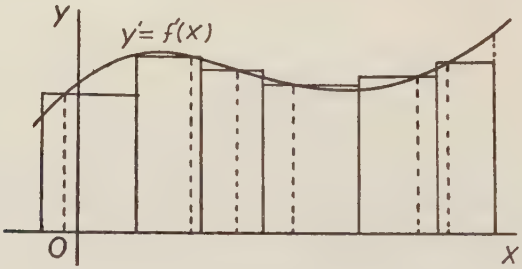


Fig. 20

sum of rectangles constructed in this way is likely to be nearer the area in question than when each is the height of the left hand ordinate, and the limit of the sum as the width of the rectangle approaches zero is really the required area (Fig. 20). Since differentials are used only in a temporary stage of a problem, and only the limit of their sum enters into final results, it is immaterial, when we write $dy = y'dx$, whether we regard y′ as the derivative of y with respect to x at the left end of the interval covered by dx or at some other place on the interval.

We will now illustrate the method of analysis by which this idea of limit of a sum of differentials is used to solve practical problems.

First, let us determine what is the moment of the force necessary to overcome friction which is distributed uniformly over one end of a right circular cylinder and acts to resist rotation of the cylinder about its axis, as in a "step bearing." Suppose the radius of the cylinder to

be a inches, and the force of friction to be p pounds per square inch of area. Consider a circle of radius a and center at O, which will be regarded as the axis of the cylinder. Let this circle be divided up (Fig. 21) by smaller circles with center at O, and let Δr denote the difference of radii of any two successive circles. The area between these two circles is $2\pi r_1 \Delta r$, where r_1 is a certain intermediate radius. The total friction force on this area is $2\pi r_1 p \Delta r$, and the moment of this force about o is

$$\Delta M = 2\pi r_1 r_2 p \Delta r, \tag{1}$$

where r_2 is also an intermediate radius. Now the sum of all such quantities ΔM is exactly the moment required, but we do not know the values of r_1 and r_2, so this sum cannot be found directly. But r_1 and r_2 lie

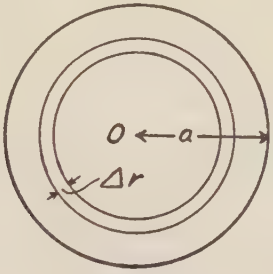


Fig. 21

between the radii of the two circles, and if $r^2 = r_1 r_2$, then r is also an intermediate radius, and if we write

$$dM = 2\pi r^2 p dr \tag{2}$$

and take the limit of the sum of such expressions, we have without finding the value of r, a definite integral which gives the value of the required moment.

The moment is

$$M = \int_0^a 2\pi p r^2 dr = \frac{2}{3}\pi p a^3.$$

Now this problem may be solved without the notion of differential, for from (1) we have

$$\frac{\Delta M}{\Delta r} = 2\pi r_1 r_2 p,$$

and as Δr approaches zero these intermediate radii r_1 and r_2 approach coincidence at r, therefore the derivative is

$$\frac{dM}{dr} = 2\pi pr^2,$$

and

$$M = \int_0^a 2\pi pr^2 dr = \frac{2}{3}\pi pa^3.$$

For most problems of this character, however, the use of the differential is preferred, the differential dM being regarded as an approximation for the increment ΔM, and the limit of the sum of the dM's is exactly the sum of the ΔM's. In practice we do not write equation (1), but write at once equation (2), realizing that any intermediate radius, as r_1 or r_2, may be replaced by r.

Now let us determine how long it will take to empty a rectangular tank of water through an orifice at the bottom, given that the rate of flow through the orifice in cubic inches per minute is equal to $k\sqrt{h}$, where h is the depth of water above the orifice, and k is a constant. Let the width and breadth of the tank be a and b, and let h_1 be the initial depth of water. The time required to lower the surface a distance dh is

$$dt = \frac{abdh}{k\sqrt{h}},$$

that is, the volume removed divided by the volume removed per minute. We have

$$t = \frac{ab}{k}\int_0^{h_1} \frac{dh}{\sqrt{h}} = \frac{2ab}{k}\sqrt{h_1}.$$

The quantities a, b and h_1 being in inches, t is given in minutes. Here, in the expression for dt, to find the exact interval of time we should have to find the average value of $k\sqrt{h}$ while h changes through the distance dh, but in using the differential and taking the limit of the sum, i.e., taking the definite integral, we are not required to know this average value.

EXERCISES XV

1. The load per foot run on a certain beam at distance x feet from one end of the beam is given by $L = 100x$, where x is in feet. If the

beam is 10 feet long what is the total load? (Hint: Find the differential dP of load corresponding to length dx of beam, then integrate).

2. Find the moment of the load in example 1 about the end from which x is measured. Show that if the beam is supported at the two ends, two-thirds of the load is supported at the end where $x = 10$.

3. Understanding that the shearing force at any point on a beam is the resultant of all vertical forces on one side of that point, including load and supporting forces, find the shearing force at the middle of the beam of examples 1 and 2. Find the shearing force at any distance x, as a function of x.

4. A certain coiled wire spring expands one inch for each 10 pounds of force applied. How much work is done in stretching the spring 8 inches? (See example 3, XIV.)

5. A flat collar bearing on a shaft takes up a thrust of 100 lbs. The inner radius of the bearing is 3 inches, the outer radius 5 inches. Considering the pressure uniform on the surface of the bearing and the coefficient of friction .05, find the moment of the friction force on the bearing.

6. A tank in the form of a cylinder with axis vertical is to be emptied through an orifice of one square inch. If the velocity of flow in feet per second is $5\sqrt{h}$, where h is the depth in feet of water above the orifice, find how long it will take to lower the surface of water 4 feet from a depth of 10 feet, the radius of the tank being 6 feet.

7. Assuming that the pressure in pounds per square foot at a point on a surface under water is 62.5 h, where h is the depth of water in feet above the point, find the total pressure on a vertical wall 6 feet high and 10 feet wide, just submerged. Find the total pressure on the lower half of this wall.

8. Find the moment about the upper edge of the wall in example 7 of the pressure due to the water. Where should a concentrated force be applied to balance this pressure?

9. A funnel is in the form of a cone, 10 inches in diameter, 12 inches altitude. Water flows out at the rate of $10\sqrt{h}$ cubic inches per second, where h is depth of water in inches. Find the time required to empty the funnel from a depth of 8 inches.

CHAPTER V

SOLUTION OF EQUATIONS

Art. 19. Algebraic Equations, Exact Roots

Let P(x) denote a polynomial of the nth degree in x,

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

where n is a positive integer and the a's are independent of x.

Remainder theorem: If P(x) is divided by x − r the remainder is P(r).

Proof: Letting $P(x) = Q(x)(x - r) + R$, where Q(x) is the quotient and R the remainder, we have, putting $x = r$, $P(r) = R$.

The following important theorem is an immediate corollary of the remainder theorem, since to say that r is a root of the equation $P(x) = 0$ is to say that the remainder, P(r), is zero.

Factor theorem: If r is a root of the equation $P(x) = 0$, then P(x) is exactly divisible by x − r.

From the factor theorem above we see that if r is a root of equation

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0. \qquad (1)$$

then x − r is a factor of the left hand member, and so if $r_1, r_2 \ldots r_n$ are roots $(x - r_1)(x - r_2)(x - r_3) \ldots (x - r_n)$ is a factor of the left member of the equation. There cannot be more than n roots, for that gives a polynomial of degree higher than n.

Assuming that an equation of the nth degree has at least one root, we see that it must have n roots, and therefore must have the structure.

$$a_0(x - r_1)(x - r_2) \ldots (x - r_n) = 0. \qquad (2)$$

Multiplying out, we obtain

$$a_0x^n + a_0p_1x^{n-1} + a_0p_2x^{n-2} + \ldots a_0p_{n-1}x + a_0p_n = 0 \qquad (3)$$

where p_1 is minus the sum of roots, p_2 is the sum of all the products two at a time, and so on, p_k being the sum of all the products of roots k at a time, negative if k is odd.

It is worth while here to consider some special types of polynomials.

First let $a_o = 1$ and let the roots $r_1, r_2, \ldots, r_n$ have respectively the values of $0, 1, 2, \ldots, n-1$, and let us consider the value of the polynomial when x is a positive integer, say m. The left member of (2) now becomes the product of n successive integers, and is assigned a special notation.

$$ {}_mP_n = m(m-1)(m-2)\ldots(m-n+1). \tag{4}$$

This quantity is zero if m is less than n, but if $m = n$ we have, with a new notation,

$$ {}_nP_n = n(n-1)(n-2)\ldots 2\cdot 1 = n! \tag{5}$$

This quantity is called **factorial** n, and is used in a great variety of applications. It is evident that if m is greater than n we have

$$ {}_mP_n = \frac{m!}{(m-n)!} = \frac{m(m-1)\ldots 2\cdot 1}{(m-n)(m-n-1)\ldots 2\cdot 1}. \tag{6}$$

An important use of these notations may be shown by an example. The number of different ways that the 5 positions on a basket ball team can be filled from among 20 athletes may be shown as follows: The first position may be filled by any one of the 20 men. For each way that the first is filled, there are 19 ways to fill the second position, so that the first two may be filled in $20\cdot 19$ ways, and for each of these ways the next position may be filled in 18 ways, and so on, the required number being the product of five factors, 20, 19, 18, 17, 16. This is the number of **permutations** of 20 things taken 5 at a time, and may be represented,

$$ {}_{20}P_5 = \frac{20\cdot 19\cdot 18\cdot 17\ldots 3\cdot 2\cdot 1}{15\cdot 14\cdot 13\ldots 3\cdot 2\cdot 1} = \frac{20!}{15!}. $$

In general, the number of **permutations of n things taken k at a** time is

$$ {}_nP_k = \frac{n!}{(n-k)!}. \tag{7}$$

In case $k = 1$, we have ${}_nP_1 = \frac{n!}{(n-1)!} = n$, obviously the number of ways that one may be selected from n things. If $k = n$, we have as an important special case ${}_nP_n = \frac{n!}{0!}$. Now 0! must be assigned the

value 1, otherwise the identity $\frac{n!}{(n-1)!} = n$ would fail when $n = 1$; so that ${}_nP_n = n!$ as in (5), and is easily shown to be the number of different ways that n things may be arranged or permuted.

Now if we wish to find how many different committees of five may be appointed from among these athletes, we observe that each committee could fill the 5 positions on the basket ball team in 5! different ways, so if we divide the number ${}_{20}P_5$ by 5!, we have the number of committees. This is the number of **combinations** of 20 things taken 5 at a time, and is represented

$$ {}_{20}C_5 = \frac{20!}{5! \cdot 15!}. $$

In general the number of **combinations of** n **things taken** k **at a time** is

$$ {}_nC_k = \frac{n!}{k!\,(n-k)!}. \tag{8} $$

It will be observed that a committee of 15 could be appointed in as many different ways as the committee of 5, since every time 5 are selected, 15 are left. In general ${}_nC_k = {}_nC_{n-k}$.

As a second special type of polynomial, let $a_o = 1$ and let all the roots be equal, say $r_1 = r_2 = \ldots = r_n = r$, then the left member of (2) becomes $(x - r)^n$, while in (3) we have $p_1 = -nr$, $p_n = \pm r^n$, and in general $p_k = \pm {}_nC_k\, r^k$, the positive sign pertaining to even powers of r. The relation becomes

$$ (x-r)^n = x^n - nrx^{n-1} + \frac{n!}{2!\,(n-2)!} r^2 x^{n-2} \ldots \pm r^n. \tag{9} $$

This is known as the **Binomial Theorem for positive integral powers,** and by use of this equation we are able to write at once any required term in any positive integral power of a binomial. For example the **kth term** in $(a - b)^n$ involves b^{k-1}, and may be written,

$$ (-1)^{k-1} \frac{n!}{(k-1)!\,(n-k+1)!} a^{n-k+1} b^{k-1}. \tag{10} $$

Returning now to the polynomial equation (3) we have the important theorem:

If the coefficients in (3) **are all integers, and if** p **and** q **are integers having no common factor, then if** $\frac{p}{q}$ **is a root of** (3) q **is a factor of** a_o.

Proof: Suppose $x = \frac{p}{q}$ in (3), then we have

$$\frac{a_o p^n}{q^n} + \frac{a_o p_1 p^{n-1}}{q^{n-1}} + \ldots + a_o p_n = 0,$$

or, multiplying by q^{n-1} and transposing all but the first term,

$$\frac{a_o p^n}{q} = a_o p_1 p^{n-1} + a_o p_2 p^{n-2} q + \ldots + a_o p_n q^{n-1}.$$

Here the right member is an integer, hence also the left member, and therefore q must be a factor of a_o.

As a result of this theorem we see that in an equation in the form

$$x^n + p_1 x^{n-1} + p_2 x^{n-2} + \ldots + p_n = 0, \qquad (11)$$

in which the p's are all integers, the rational roots are integers, factors of p_n. Furthermore, if some of the p's are not integers, but are rational fractions, we may form a new equation, say

$$y^n + q_1 y^{n-1} + q_2 y^{n-2} + \ldots + q_n = 0, \qquad (12)$$

where the q's are integers and such that the roots of (12) are multiples of the roots (11). To **multiply the roots** of (11) by m, that is to determine the q's in (12) so that the values of y are the values of x multiplied by m, we put $y = mx$ in (12), then divide by m^n and obtain

$$x^n + \frac{q_1}{m} x^{n-1} + \frac{q_2}{m^2} x^{n-2} + \ldots + \frac{q_n}{m^n} = 0.$$

Comparing this with (11) we see that

$$q_k = m^k \cdot p_k \ (k = 1, 2, \ldots n). \qquad (13)$$

This result may be seen otherwise, since p_k is the sum of products of roots taken k at a time, and multiplying each root by m multiplies p_k by m^k. If the p's are rational m may always be chosen so that the q's are integers.

Given an equation in the form (1) in which the a's are known as rational numbers, we may always **find the rational roots** by trial. For, by dividing by a_o we reach the form (11) and, multiplying the roots by m if necessary, we obtain an equation (12) in which the rational

roots are integers. Since the product of the roots of (12) is the integer q_n, we try each of the factors of q_n to see if it is a root. By the factor theorem we know that if r is a root then $x - r$ is a factor, so that the trial is usually made most conveniently by **synthetic division.** This operation is an instance of the use of detached coefficients. In multiplying or dividing one polynomial by another we may avoid writing the several powers of variable after the coefficients if we are careful to indicate by the positions and order of the coefficients the power of variable to which they belong. In the present instance we are concerned with the division of a polynomial by a binomial of the form $x - r$. The elementary method of division may be illustrated as follows:

$$
\begin{array}{rl}
4x^3 - 3x^2 + 7x - 6 & \big| \; x - 3 \\
\underline{4x^3 - 12x^2} & \big| \; \underline{4x^2 + 9x + 34} \\
9x^2 + 7x & \\
\underline{9x^2 - 27x} & \\
34x - 6 & \\
\underline{34x - 102} & \\
96 &
\end{array}
$$

This work may be abridged by omitting the powers of x and by avoiding repetition.

$$
\begin{array}{rl}
4 - 3 \quad 7 - 6 & \big| \; 1 - 3 \\
\underline{- 12} & \big| \; \underline{4 \quad 9 \quad 34} \\
9 & \\
\underline{-27} & \\
34 & \\
\underline{-102} & \\
96 &
\end{array}
$$

Now we may arrange the work more compactly, and by repeating the first coefficient, 4, we have three numbers which determine the quotient, and the remainder, 96, arranged in proper order.

$$
\begin{array}{llll|l}
4 & -3 & 7 & -6 & 1 \; -3 \\
 & -12 & -27 & -102 & \\
\hline
4 & 9 & 34 & 96 &
\end{array}
$$

Since the first term of the divisor is understood to be x, this may be

omitted, and by changing the sign of the second term we add instead of subtracting to find remainders. The final form is:

4	−3	7	−6	|3
	12	27	102	
4	9	34	96	

As another illustration, we find the result of dividing $x^5 - 6x^3 + 3x^2 - 8$ by $x + 2$. We have by synthetic division,

1	0	−6	3	0	−8	|−2
	−2	4	4	−14	28	
1	−2	−2	7	−14	20	

where the quotient is $x^4 - 2x^3 - 2x^2 + 7x - 14$ and the remainder is 20.

When by trial we find $x - r$ to be an exact divisor we turn attention to the **depressed equation,** found by setting the quotient equal to zero. The number of trials for positive integral roots is limited not only by the number of factors of q_n, but also by the fact that we are able to set a **bound for the positive roots,** and factors of q_n greater than this bound need not be tried. **If in a trial division by** $x - b$ **the coefficients in the quotient and the remainder are all zero or positive, then no root is greater than** b. In practice, then, we try for positive roots first, depressing the equation if roots are found. To find the **negative roots** we multiply the roots by −1, then try for positive roots in the new equation. By (13) above we see that changing the signs of the roots is equivalent to changing the signs of alternate powers of the unknown.

When all the rational roots have been found and the corresponding factors removed, the depressed equation may be solved exactly by radicals if it is not above the fourth degree. In Art. 23 we discuss approximate methods of solving equations, and in most practical cases those methods are preferable to solution by radicals for irrational roots, except in case of **quadratic equations** and **binomial equations.** Accordingly we confine attention here to these two types.

If the given equation is of second degree, it may be put in the general quadratic form

$$ax^2 + bx + c = 0$$

and the roots may be found at once by one of two methods.

First, if the quadratic expression can be factored by inspection we have the form

$$(hx + k)(sx + t) = 0$$

and the roots are

$$x = -\frac{k}{h}, -\frac{t}{s}.$$

Second, if the factors are not readily found by inspection, we may write the roots at once by a formula derived as follows: We complete the square in x, first multiplying the equation through by 4a to avoid fractions.

$$4a^2x^2 + 4abx + 4ac = 0,$$

$$4a^2x^2 + 4abx + b^2 = b^2 - 4ac,$$

$$2ax + b = \pm\sqrt{b^2 - 4ac},$$

whence the formula

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a}.$$

The direct use of this formula is decidedly advantageous in case a, b, and c are expressions involving arbitrary constants or additional variables. The quantity $b^2 - 4ac$ is known as the **discriminant** of the quadratic equation, since the roots are real and unequal, real and equal, or conjugate imaginaries, according as $b^2 - 4ac$ is positive, zero, or negative, provided only that a, b, and c are real.

If the equation is in binomial form,

$$x^k - q = 0,$$

then the roots are the kth roots of q, and may be indicated by a radical. If k is even and q positive there will be one positive real root and one negative real root; if k is even and q negative all roots are imaginary; if k is odd there is one real root, agreeing in sign with q. The character of the roots of a binomial equation is discussed further in Art. 50.

Example: Solve $6x^4 - 11x^3 + 9x^2 + 4x - 2 = 0$.

Divide by 6: $x^4 - \frac{11x^3}{6} + \frac{3x^2}{2} + \frac{2x}{3} - \frac{1}{3} = 0.$

Put $y = 6x$: $y^4 - 11y^3 + 54y^2 + 144y - 432 = 0.$

Try 1;	1	−11	54	144	−432	1	f(1) = −244
		1	−10	44	188		
	1	−10	44	188	−244		
Try 2;	1	−11	54	144	−432	2	
		2	−18	72	432		
	1	− 9	36	216	0		

One root is 2, and the depressed equation is

$$y^3 - 9y^2 + 36y + 216 = 0$$

Again trying 2, then 3, 4, 6, 8, we find no root, and trying 9

1	−9	36	216	9
	9	0	324	
1	0	36	540	

All remainders are positive or zero, so that no real root, rational or irrational, is as great as 9.

Changing sign of roots, or putting $z = -y$.

$$z^3 + 9z^2 + 36z - 216 = 0$$

By trial 1 and 2 are not roots, but for 3 we find

1	9	36	−216	3
	3	36	216	
1	12	72	0	

so that 3 is a root, and we have yet to solve $z^2 + 12z + 72 = 0$.

$$z = \frac{-12 \pm \sqrt{144 - 288}}{2} = -6(1 \pm \sqrt{-1}).$$

The four roots of the original equation are $\frac{1}{3}$, $-\frac{1}{2}$, $1 + \sqrt{-1}$ and $1 - \sqrt{-1}$.

EXERCISES XVI

1. How many arrangements can be made of 12 things taken five at a time?

2. How many signals may be made with five flags, each signal to consist of three flags in a row.

3. As a bonus for extending their line a railroad company is allowed to select six sections (square miles) of land from a township six miles square. If no two sections selected may fall in the same row of sections, north and south or east and west, in how many ways may the selection be made? How many ways if the choice is unrestricted?

4. In how many ways may a detail of seven men be selected from 25 men.

5. What is the coefficient of a^5b^8 in the expansion of $(a + b)^{13}$? If 13 apples and 13 beets are paired off, one apple with each beet, how many ways may we select 5 apples and 8 beets, taking one from each pair?

6. Write the 12th term of expansion of $\left(\frac{1}{x} - x^{\frac{3}{2}}\right)^{17}$.

7. Solve the following quadratic equations:

(a) $x^2 - 6x + 5 = 0$. (c) $3x^2 + 7x - 2 = 0$.

(b) $x^2 - 13x - 30 = 0$. (d) $x^2 - kx + 4 = 0$.

8. Write the equation whose roots are 3 and -4; 2 and 8; a and $-a$; $a + \sqrt{b}$ and $a - \sqrt{b}$.

9. Solve for x; $x^2 + 4xy + 4y^2 - x + 6 = 0$.

For what value of y will the two values of x be equal? When will x be real? When imaginary?

10. An equation is to have the form of $x^2 + 2ax - a = 0$. If one root is twice the other, what must be the value of a? If one root exceeds the other by 5 what must be the value of a?

11. Construct the equation whose roots are:

(a) 2, 5, and -3; (b) 1, 3, and -4;

(c) 2, 3, -1, -5; (d) $2, -1, 3 + \sqrt{-2}, 3 - \sqrt{-2}$.

12. Find the rational roots of the following equations, and complete the solution when possible.

(a) $x^3 - 7x^2 + 14x - 8 = 0$. (b) $x^4 - 2x^3 - 12x^2 - 13x - 10 = 0$.

(c) $6x^4 + 5x^3 + 5x - 6 = 0$. (d) $5x^3 + \frac{7}{6}x^2 + \frac{27}{5} = 0$.

(e) $2x^5 - 3x^4 - 5x^3 + 4x^2 - 6x - 10 = 0$.

Art. 20. Determinants and Elimination

In the solution of a problem it may be required to find the value of one or more unknown quantities. If the problem is a definite or determinate one, there must be information at hand by which we may write as many independent equations involving these unknown quantities as there are unknown quantities. If in the statement of these

equations it is found necessary or expedient to introduce a certain number of auxiliary unknowns, an equal number of additional equations must be written.

In general, to find n unknowns we require n independent equations i.e., n equations no one of which is a consequence of the others. By transforming and combining these equations in various ways to form new equations, we may arrive at a smaller set, n − r equations in n − r unknowns. The r unknowns that do not appear are said to be **eliminated.** To find the value of an unknown, it is necessary to carry this process of elimination on until all other unknowns are eliminated, leaving one equation in the required unknown. This elimination is not always possible, and when possible may be very laborious. It is the purpose of the present article to show a practical and systematic method of elimination when the equations may be reduced to rational algebraic equations.

1. **Determinant Notation:** Consider a set of three linear equations in three unknown quantities, x, y and z.

$$a_1x + b_1y + c_1z = d_1,$$

$$a_2x + b_2y + c_2z = d_2,$$

$$a_3x + b_3y + c_3z = d_3.$$

Elimination of y and z (by any method) gives

$$x = \frac{d_1b_2c_3 + b_1c_2d_3 + c_1d_2b_3 - c_1b_2d_3 - d_1c_2b_3 - b_1d_2c_3}{a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 - a_1c_2b_3 - b_1a_2c_3}.$$

It will be noticed that the six terms of the denominator each involve three different letters and three different subscripts; and they are the only terms of this kind that can be formed from the coefficients of the three equations, since, the subscripts being in natural order, they are the six permutations of the three letters. It may be observed that the positive terms have the letters in natural order, or an even number of interchanges or **inversions** of order, while the negative terms have an odd number of inversions of order. The numerator is like the denominator, with d's in place of a's. Such expressions are called **determinants.**

The notation for a determinant is a square array of elements, in rows and columns, bordered on right and left by vertical lines. Thus, a determinant of third order has three rows and three columns, as follows:

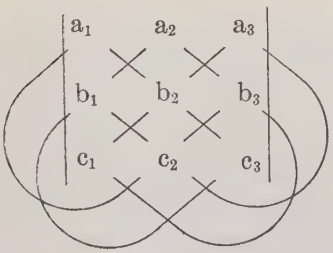
$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

The value of a determinant is the sum (algebraic) of all products that can be formed by taking one element from each row and column, the sign of each product (**term**) to be + if the number of inversions is even, and − if the number of inversions is odd. For the purpose of determining the number of inversions in a given term the term may be written, as $a_2\ c_1\ b_3$, and the cases of "unnatural" order may be counted. Thus, in the term written, the inversions are: c before b, 2 before 1; it being understood that the elements are lettered in the order of the rows and numbered in the order of the columns. Change of the order of elements in a term can change the number of inversions by an even number only, so that the criterion always yields a definite sign for the term.

2. **Expansion of Determinants:** The foregoing definition leads to simple rules for finding the value of a determinant. For the determinant of the second order we have:

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1,$$

and for the determinant of the third order we have:



$$= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1c_2b_3 - a_2b_1c_3 - a_3b_2c_1.$$

The lines drawn through the elements show the manner of selecting terms, and those that slope one way give + signs, while those that slope the other wave give − signs.

Example:

$$\begin{vmatrix} 2 & 1 & -4 \\ 3 & -2 & 1 \\ 3 & 4 & -5 \end{vmatrix} = 2\cdot(-2)\cdot(-5) + 1\cdot1\cdot3 + (-4)\cdot4\cdot3 - (-4)\cdot(-2)\cdot3 - 1\cdot3\cdot(-5) - 2\cdot1\cdot4 = 20 + 3 - 48 - 24 + 15 - 8 = -42.$$

From the definition of a determinant it will be evident also that any determinant may be expressed in terms of determinants of lower order, called **minor** determinants. The minor corresponding to a given element is found by suppressing the row and column in which the element occurs. Thus, in the third ordered determinant above the minors which correspond respectively to a_1 and b_3 are:

$$\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}; \qquad \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}.$$

Any determinant may be expressed as the sum of the elements of any row or column, each multiplied by its corresponding minor determinant with the proper algebraic sign prefixed. The signs to go with the different elements alternate in rows and in columns, the first element in the first column being +.

$$\begin{vmatrix} a_2 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

$$= -a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}$$

$$= c_1 \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Determinants of higher order than the third may thus be expressed in terms of third ordered determinants by successive reductions, and

their values found by evaluating these third ordered determinants. Since a determinant of the nth order has in general n! terms, the process of expancing may be very laborious. A fifth ordered determinant clearly leads to 20 third ordered determinants or 60 second ordered determinants.

3. **Certain properties of determinants.** While the method indicated above for expanding a determinant is always effective, it may often yield results more quickly if combined with the use of certain general properties of determinants. These properties also aid in recognizing relations that may exist among different determinants.

(a) **The value of a determinant is not altered if columns and rows are interchanged, the first column replacing the first row, etc.**

(b) **If two rows (or columns) are interchanged the sign of the determinant is changed.**

For this changes the number of inversions in each term of the expanded determinant by an odd number.

(c) **If the elements of a row (or column) contain a common factor, this factor is a factor of the determinant.**

For one and only one element of this row (or column) is found as a factor in each term of the expanded determinant.

(d) **If the elements of one row (or column) are proportional to the corresponding elements of another row (or column), the determinant is zero.**

For, we may remove a factor from one row (or column) and leave two rows (or columns) equal. Then expressing the determinant first in terms of the elements of one row (or column) and their minors, then the other row and their minors, the two expressions differ only in sign, yet must be equal, hence must be zero.

(e) **If the elements of one row (or column) are altered by adding or subtracting multiples of the corresponding elements of another row (or column) the value of the determinant is not changed.**

For example:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 + ma_2 & a_2 & a_3 \\ b_1 + mb_2 & b_2 & b_3 \\ c_1 + mc_2 & c_2 & c_3 \end{vmatrix}$$

$$= (a_1 + ma_2)\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - (b_1 + mb_2)\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + (c_1 + mc_2)\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} ma_2 & a_2 & a_3 \\ mb_2 & b_2 & b_3 \\ mc_2 & c_2 & c_3 \end{vmatrix},$$

since this last determinant is zero, by property (d).

The property (e) is used to replace as many elements as possible by zero, or by smaller numbers, so that the labor of evaluating is reduced. We illustrate by evaluating a determinant of the fourth order.

$$\begin{vmatrix} 22 & 3 & 2 & -2 \\ 32 & 6 & 7 & 3 \\ 14 & 2 & 4 & 1 \\ 24 & 4 & -3 & 5 \end{vmatrix} = \begin{vmatrix} 4 & 3 & 2 & -2 \\ -4 & 6 & 7 & 3 \\ 2 & 2 & 4 & 1 \\ 0 & 4 & -3 & 5 \end{vmatrix} = \begin{vmatrix} 0 & -1 & -6 & -4 \\ 0 & 9 & 9 & 1 \\ 2 & 2 & 4 & 1 \\ 0 & 4 & -3 & 5 \end{vmatrix}$$

$$= -2 \begin{vmatrix} 1 & 6 & 4 \\ 9 & 9 & 1 \\ 4 & -3 & 5 \end{vmatrix} = -2 \begin{vmatrix} -35 & 5 & 4 \\ 0 & 0 & 1 \\ -41 & -7 & 5 \end{vmatrix} = 2 \cdot 5 \begin{vmatrix} -7 & 1 \\ -41 & -7 \end{vmatrix} = 900.$$

The first step is made by diminishing the first column by 6 times the second; the second step by adding the first row to the second then subtracting twice the third from the first. The third step expresses it in terms of elements of the first column and their minors, the fourth is by taking the first column from the second, then nine times the third from the first. The fifth step expresses the third ordered determinant in terms of elements of the second row and their minors, and takes out the factor 5 from the first row.

4. **Uses of determinants.** The following examples serve to illustrate the more elementary and more important uses of determinants.

(a) **To solve a set of n simultaneous linear equations in n unknowns.** Example, n = 3.

$$a_1x + a_2y + a_3z = a_4,$$

$$b_1x + b_2y + b_3z = b_4,$$

$$c_1x + c_2y + c_3z = c_4.$$

$$x = \frac{\begin{vmatrix} a_4 & a_2 & a_3 \\ b_4 & b_2 & b_3 \\ c_4 & c_2 & c_3 \end{vmatrix}}{\Delta}, \quad y = \frac{\begin{vmatrix} a_1 & a_4 & a_3 \\ b_1 & b_4 & b_3 \\ c_1 & c_4 & c_3 \end{vmatrix}}{\Delta}, \quad z = \frac{\begin{vmatrix} a_1 & a_2 & a_4 \\ b_1 & b_2 & b_4 \\ c_1 & c_2 & c_4 \end{vmatrix}}{\Delta},$$

where Δ is the determinant of the coefficients,

$$\Delta = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

The proof consists in substituting the expressions for x, y and z in each equation transposing the constant term, and multiplying by Δ, when the left member of each equation will be a fourth ordered determinant with two rows equal, and is therefore zero.

In general, the denominator, for each of the n unknowns, is the determinant of the coefficients, and the numerator is that same determinant with the coefficients of the unknown in question replaced by the constant terms.

Two special cases must be noted. First, if Δ, the denominator, is zero, while the other determinants are not zero, then the equations are not consistent, there being no finite values of the unknown quantities that can satisfy the equations. Second, it is shown as an example in (b) below, that if all the determinants are zero the equations are not independent, and are therefore not sufficient to determine the unknown quantities.

(b) **To investigate the consistency of a redundant set of equations.** A set of equations is said to be redundant in case there are more equations than there are unknown quantities. Since n independent equations determine n unknown quantities, the remaining equations of the set will be satisfied by the values so determined only if they are specially related to the n equations. The relations will be made clear by an example.

Let it be required to see whether the third equation in (a) above can be formed from the other two, i.e., we are to find p and q so that p times the first equation added to q times the second gives the third equation. We should have

$$\begin{aligned} a_1p + b_1q &= c_1, \\ a_2p + b_2q &= c_2, \\ a_3p + b_3q &= c_3, \\ a_4p + b_4q &= c_4. \end{aligned}$$

Here are four equations to determine the two unknowns, p and q, a redundant set. Solve the first two equations,

$$q = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \quad q = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}.$$

Substituting these values in the third equation, and making obvious transformations,

$$a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = 0,$$

$$\text{or} \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0.$$

This equation expresses the condition that there exist values of p and q that satisfy the first three equations. A similar condition may be expressed for any three of the four equations. We see then by property (a) of Section 3 that if the four determinants in (a) above are zero, the three equations are not independent.

With regard to a redundant set of equations, it may be shown by reasoning as above, that n + r linear equations in n unknowns are consistent only when every determinant of order n + 1, formed from n + 1 of the equations, is zero. The case of most frequent interest is when r = 1, then only one determinant is involved.

Similarly the equations

$$a_1x + b_1y + c_1z = 0,$$

$$a_2x + b_2y + c_2z = 0,$$

$$a_3x + b_3y + c_3z = 0,$$

have solutions other than x = y = z = 0, only in case the determinant is zero. In this case only the ratios, $\frac{x}{z}$ and $\frac{y}{z}$ may be determined. Any two of the three equations may be solved for these values.

(c) **To express the condition that two algebraic equations in one unknown have a common root.**

Example:

$$x^2 + 3x - a = 0,$$

$$x^3 - ax^2 + 2x - 5 = 0.$$

Write these equations and others found by multiplying by x, as follows:

$$\begin{aligned} x^2 + 3x - a &= 0, \\ x^3 + 3x^2 - ax &= 0, \\ x^4 + 3x^3 - ax^2 &= 0, \\ x^3 - ax^2 + 2x - 5 &= 0, \\ x^4 - ax^3 + 2x^2 - 5x &= 0. \end{aligned}$$

Any value of x that satisfies the two given equations will satisfy all five of these. Regarding x, x^2, x^3, and x^4 as unknowns, we see from (b) that the five equations can be satisfied only in case

$$\begin{vmatrix} 0 & 0 & 1 & 3 & -a \\ 0 & 1 & 3 & -a & 0 \\ 1 & 3 & -a & 0 & 0 \\ 0 & 1 & -a & 2 & -5 \\ 1 & -a & 2 & -5 & 0 \end{vmatrix} = 0.$$

If a is chosen so this is true, the equations have a common root. The determinant is known as the **eliminant** of the two given equations.

(d) **To eliminate one variable from two simultaneous equations.**

Example:

$$x^2y - 3xy^2 + 2y - 1 = 0,$$

$$xy - y^2 + 2 = 0.$$

Proceeding as before

$$\begin{aligned} x^2y - 3xy^2 + (2y - 1) &= 0, \\ xy - (y^2 - 2) &= 0, \\ x^2y - (y^2 - 2)x &= 0. \end{aligned}$$

In order that x and x^2 may satisfy these three equations we must have

$$\begin{vmatrix} y & -3y^2 & (2y-1) \\ 0 & y & -(y^2-2) \\ y & -(y^2-2) & 0 \end{vmatrix} = 0.$$

This gives an equation for y, i.e., x is eliminated. The determinant may be called the **x-eliminant** of the two equations.

The method used here and in (c) is known as **Sylvester's method of elimination,** and consists in multiplying the equations by powers of the variable to be eliminated, until the coefficients and terms free from the variable form a square array. If the given equations are of degree m and n, we require in all m + n equations.

EXERCISES XVII

1. Evaluate the following determinants:

(a) $\begin{vmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \\ 6 & 1 & 2 \end{vmatrix}$; (b) $\begin{vmatrix} 1 & 0 & -3 \\ 0 & 2 & 4 \\ -2 & 0 & -8 \end{vmatrix}$; (c) $\begin{vmatrix} x & y & 1 \\ 4 & 3 & 1 \\ -2 & 7 & 1 \end{vmatrix}$;

(d) $\begin{vmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 0 & 0 \\ 6 & 1 & 2 & 0 \\ 0 & 2 & 0 & -1 \end{vmatrix}$; (e) $\begin{vmatrix} 0 & a & b & c \\ 0 & 1 & d & e \\ a & b & c & 0 \\ 1 & d & e & 0 \end{vmatrix}$.

2. The following identical relations should be proved, if possible, by the aid of properties in Section 3, without expanding the determinants

(a) $\begin{vmatrix} 2 & 1 & 6 \\ 5 & 2 & 15 \\ 3 & 7 & 9 \end{vmatrix} = 0$; (b) $\begin{vmatrix} 5 & -3 & 1 \\ 6 & 1 & -2 \\ -10 & 6 & -2 \end{vmatrix} = 0$;

(c) $\begin{vmatrix} 4 & 8 \\ 6 & 3 \end{vmatrix} = -\begin{vmatrix} 8 & 3 \\ 4 & 6 \end{vmatrix}$; (d) $\begin{vmatrix} 2 & -6 \\ 8 & 9 \end{vmatrix} = \begin{vmatrix} 3 & -4 \\ 12 & 6 \end{vmatrix}$;

(e) $\begin{vmatrix} 3 & -2 & 1 \\ -1 & 2 & -3 \\ 7 & -3 & 5 \end{vmatrix} = \begin{vmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \\ 7 & 3 & 5 \end{vmatrix}$; (f) $\begin{vmatrix} 3 & 5 & 7 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 3 & 5 & 1 \\ 7 & 4 & 1 & 1 \end{vmatrix} = 0.$

3. Expand the following determinants:

(a) $\begin{vmatrix} 48 & 5 & 9 \\ 36 & 6 & 7 \\ 72 & 8 & 14 \end{vmatrix}$; (b) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{vmatrix}$; (c) $\begin{vmatrix} 1 & 4 & 5 \\ 3 & 11 & 16 \\ 7 & 30 & 36 \end{vmatrix}$;

(d) $\begin{vmatrix} a & 1 & 0 & 0 & 0 \\ 1 & a & 0 & 0 & 0 \\ 0 & 0 & b & 1 & c \\ x & y & 1 & 1 & 0 \\ y & x & c & 0 & 0 \end{vmatrix}$; (e) $\begin{vmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 6 & 2 & 1 & 0 \\ 0 & 0 & 8 & -2 & 1 \\ 1 & 0 & 0 & 1 & 2 \end{vmatrix}$.

4. Solve the following by determinants:

(a) $3x + 2y = 7,$
$4x - y = 2;$

(b) $ax - 8y = 1,$
$2x - ay = \frac{1}{2};$

(c) $\frac{1}{x} - \frac{1}{y} = 1,$
$\frac{1}{x} + \frac{1}{y} = 5;$

(d) $x + y = -1,$
$y + z = 1,$
$z + x = 4;$

(e) $2x + y + 2z = 2,$
$x - 3y + 5z = -9,$
$3x + 2y - z = 1;$

(f) $x - w = -1,$
$3x + y = 5,$
$y + 2z = 0,$
$x - 3z = 4;$

(g) $x^2 + 3xy = 4,$
$3x^2 - xy = 1;$

(h) $3x^2 + 4y^2 = 48,$
$6x^2 - y^2 = 15.$

5. Which of the following systems of equations admit of definite solutions? Find solutions when possible.

(a) $3x - 2y + z = 2,$
$x + 5y - z = 6,$
$7x + y + z = 10;$

(b) $3x + 4y = 7,$
$2x - 5y = -3,$
$x + 3y = 6;$

(c) $\begin{aligned} u + v - 3 &= 0, \\ 3u - v + 11 &= 0, \\ u - 3v + 17 &= 0, \\ 5u - 3v + 25 &= 0; \end{aligned}$ (d) $\begin{aligned} x - y + z &= 2, \\ 2x + y - 3z &= 5, \\ x + 2y - 4z &= 0. \end{aligned}$

6. In each of the following sets of equations show that one unknown may be assigned any desired value, and that solutions may then be found for the system. In each case solve for y and z in terms of x.

(a) $\begin{aligned} x - 3y &= 4, \\ 2y + z &= 7, \\ x + y + 2z &= 18; \end{aligned}$ (b) $\begin{aligned} x + y - z &= 0, \\ 3x + y + z &= 0, \\ x - y + 3z &= 0. \end{aligned}$

7. For what value of a will the following polynomials have a common factor? $x^2 - 2x + a$, $x^3 - ax - 2$.

8. Eliminate y from the following equations:

$$x^2 + 3xy - 2y^2 - 3x = 6.$$

$$xy + 4y^2 + 2y = -2.$$

9. Eliminate t from the following equations:

$$x = 1 - t^2, \qquad y = 2t + t^3.$$

Art. 21. Equations that Require Trigonometric Transformations

Equations involving unknown angles which enter the equation by means of their trigonometric functions, are called trigonometric equations. If several unknown angles are to be found an equal number of equations will be required, and the solution may involve simultaneous equations and elimination. A systematic method of attack is the following sequence of operations:

(i) Transform the equations so that each trigonometric function involves only one unknown angle.

(ii) Transform so that different multiples or fractions of the same unknown angle do not occur, and so that functions of sums of known and unknown angles do not occur.

(iii) Choose a single function of each unknown angle and express all other functions that occur in terms of these.

(iv) Solve as algebraic equations in which the unknowns are the chosen functions of the unknown angles.

(v) Having found the value of one function of each unknown angle determine the values of the angles, and test by substitution in the original equations.

The order of these operations may sometimes be altered, and frequently some of them are not necessary. Algebraic reductions are advisable when convenient at any stage of the work. The five processes (i) to (v), are each the subject of one section in the following:

1. **To separate the unknown angles:** When a trigonometric function involves more than one unknown angle, it is almost invariably as a sum or difference of angles. We are required to express the functions of the combined angles in terms of functions of the separate angles. The formulas most often useful for this purpose are:

$$\sin (x + y) = \sin x \cos y + \cos x \sin y. \tag{1}$$

$$\sin (x - y) = \sin x \cos y - \cos x \sin y. \tag{2}$$

$$\cos (x + y) = \cos x \cos y - \sin x \sin y. \tag{3}$$

$$\cos (x - y) = \cos x \cos y + \sin x \sin y. \tag{4}$$

$$\tan (x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \tan y}. \tag{5}$$

$$\cot (x \pm y) = \frac{\cot x \cot y \mp 1}{\cot y \pm \cot x}. \tag{6}$$

These formulas, established in trigonometry, are true for all values of x and y, and enable us to separate any number of angles. To illustrate:

$$\cos(x + y - u + v) = \cos[(x + y) - (u - v)]$$

$$= \cos(x + y) \cos(u - v) + \sin(x + y) \sin(u - v)$$

$$= (\cos x \cos y - \sin x \sin y) (\cos u \cos v + \sin u \sin v)$$

$$+ (\sin x \cos y + \cos x \sin y) (\sin u \cos v - \cos u \sin v).$$

2. **To standardize each unknown angle:** Having separated the unknown angles, any unknown angle may still enter in various forms; For example we may have functions of

$$x, 3x, \frac{x}{2}, x + \frac{\pi}{2}, \frac{\pi}{6} - \frac{x}{2}, \text{ etc.}$$

We may adopt a definite linear function of x as the standard and

express all functions in terms of this standard angle. The multiple and fractional angles are handled by use of the formulas:

$$\sin 2x = 2 \sin x \cos x. \tag{7}$$

$$\cos 2x = \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x. \tag{8}$$

$$\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}. \tag{9}$$

$$\cot 2x = \frac{\cot^2 x - 1}{2 \cot x}. \tag{10}$$

$$\sin^2 \frac{x}{2} = \frac{1 - \cos x}{2}. \tag{11}$$

$$\cos^2 \frac{x}{2} = \frac{1 + \cos x}{2}. \tag{12}$$

$$\tan \frac{x}{2} = \sqrt{\frac{1 - \cos x}{1 + \cos x}} = \frac{\sin x}{1 + \cos x} = \frac{1 - \cos x}{\sin x}. \tag{13}$$

The additive terms in the angles may be treated by use of formulas (1) to (6) above, by assigning suitable values to the notations x and y. It will be seen that formulae (7)–(10) together with those of the previous set enable us to reduce functions of any multiple of an angle to functions of the simple angle, while formulae (11)–(13) are successful in dealing only with fractions formed by repeating halving, i.e., fractional angles of the form $\frac{x}{2^n}$, and the radical forms make it usually inexpedient to use n greater than unity. For these reasons, it may be necessary to adopt as the standard angle a fraction of the unknown, and regard the unknown as a multiple angle. In selecting the standard angle, due regard should be paid to the fact that transformation to functions of a lower multiple or fractional angle gives expressions of higher degree algebraically, and to the desirability of keeping the algebraic degree as low as possible. In this connection, with a view to better algebraic form, for convenience in factoring, etc., we may occasionally resort to the following formulae:

$$\sin mx + \sin nx = 2 \sin \frac{1}{2}(m + n)x \cos \frac{1}{2}(m - n)x. \tag{14}$$

$$\sin mx - \sin nx = 2 \cos \frac{1}{2}(m + n)x \sin \frac{1}{2}(m - n)x. \tag{15}$$

$$\cos mx + \cos nx = 2 \cos \frac{1}{2}(m + n)x \cos \frac{1}{2}(m - n)x. \quad (16)$$

$$\cos mx - \cos nx = -2 \sin \frac{1}{2}(m + n)x \sin \frac{1}{2}(m - n)x. \quad (17)$$

3. **To transform to a single function of each standardized unknown:** To carry out the third step in the solution of our system of equations we require only the definitions of the six functions and the Pythagorean theorem on the right triangle (sum of squares of two legs equals square of hypotenuse). Consider an angle ϕ generated by revolving OP of length r, about O from the initial line OX. Let x and y be the coordinates of P in a rectangular system of which O is the origin and OX the positive direction of the x-axis. The angle ϕ determines only the ratios of r, x, and y. If $r = 1$, then $x = \cos \phi$ and $y = \sin \phi$. If $x = 1$, then $y = \tan \phi$ and $r = \sec \phi$. If $y = 1$, then $x = \cot \phi$ and $r = \csc \phi$. Thus it will be seen that by use of a right triangle in which the appropriate side is chosen equal to 1, all functions may be expressed without radicals in terms of two functions or, by use of radicals, in terms of any one function. It is often expedient to use two functions as sine and cosine, and avoid radicals until certain algebraic reductions are made, then finally express one of these in terms of the other, if necessary.

4. **To find the values of functions of the unknown angles:** We are now to deal with the n transformed equations, involving n unknown quantities, each of these quantities being a certain trigonometric function of a certain standard angle. The algebraic solution of this system of equations yields one or more values for each of the trigonometric functions involved. It may be necessary to reject some of the values thus obtained. For a tangent or cotangent any real value may be retained; for a sine or cosine, only real values from -1 to 1, inclusive, may be retained, while for secant or cosecant only values less than or equal to -1 or greater than or equal to 1 may be retained.

5. **To find the unknown angles:** It remains to determine the angles whose functions have been found. For each suitable value of a trigonometric function there is always an indefinitely large number of values of the angle. Two of these values, in general, are less than 2π, and the other values may be obtained from them by adding or subtracting integral multiples of 2π. Convenient forms of representing the general solutions are shown in the following, in which A is any

permissible value of x, preferably from $\frac{\pi}{2}$ to $-\frac{\pi}{2}$, inclusive, in the first and third expressions, and from 0 to π, inclusive, in thesecond expression.

With sin x or csc x given, $x = n\pi + (-1)^n A$. (18)

With cos x or sec x given, $x = 2n\pi \pm A$. (19)

With tan x or cot x given, $x = n\pi + A$. (20)

Having determined a value A, of x, other values are obtained by setting n successively equal to zero and the positive and negative integers.

Example 1: $\sin x + \sin 2x + \sin 3x = 0$.

There being only one unknown angle, x, the first step is unnecessary. By (14) we have,

$$\sin 3x + \sin x = 2 \sin 2x \cos x,$$

so that the equation becomes

$$\sin 2x + 2 \sin 2x \cos x = 0, \text{ or } \sin 2x (1 + 2 \cos x) = 0.$$

If $\sin 2x = 0$ then $2x = n\pi$ and $x = n\frac{\pi}{2}$.

If $1 + 2 \cos x = 0$, $\cos x = -\frac{1}{2}$, $x = 2n\pi \pm \frac{2\pi}{3}$.

Example 2: $\begin{cases} \tan(x + 2y) = \csc x - \cot x, \\ \cot(x + 2y) = \csc \frac{x}{2}. \end{cases}$

Adopting as the two unknowns x and $x + 2y$, the first step is not necessary. We observe that the second equation may be written,

$$\tan (x + 2y) = \sin \frac{x}{2},$$

so that, eliminating $x + 2y$, we have

$$\sin \frac{x}{2} = \csc x - \cot x,$$

or using (11)

$$1 - \cos x = 2 \csc^2 x - 4 \csc x \cot x + 2 \cot^2 x.$$

To avoid radicals we express all, in terms of sin x and cos x,

$$1 - \cos x = \frac{2}{\sin^2 x} - \frac{4\cos x}{\sin^2 x} + \frac{2\cos^2 x}{\sin^2 x},$$

$$(1 - \cos x)\sin^2 x = 2(1 - \cos x)^2,$$

$$(1 - \cos x)(\sin^2 x + 2\cos x - 2) = 0,$$

$$1 - \cos x = 0 \text{ or } \cos^2 x - 2\cos x + 1 = 0.$$

The three values of cos x are all coincident, equal to unity. Hence $x = 2n\pi$. To find y we have

$$\cot(x + 2y) = \csc\frac{x}{2} = \infty,$$

$$x + 2y = m\pi, \qquad 2y = m\pi - 2n\pi, \qquad y = m\frac{\pi}{2} - n\pi,$$

where m and n are any two integers, and therefore it is sufficient to write $y = m\frac{\pi}{2}$.

6. **Equations with inverse functions:** The expressions, $\sin^{-1}a$, arc sin a, anti-sine of a, inverse sine of a, angle whose sine is a, are only different expressions for one thing. The last expression is a definition of the thing, while the first and second are standard notations for the thing. This thing is an angle, and here regarded as a function of its sine, a. As shown in Section 5 this function of a has an indefinitely large number of values for each value of a. Each of the other trigonometric functions gives rise to an inverse function with analogous notation and terminology.

Equations often arise which state relations among angles, where some of the angles are inverse functions of unknown quantities. If every term in such an equation is an angle, the equation is reduced to an algebraic equation by taking a trigonometric function of both sides of the equation.

Example: Let it be required to find x from the equation

$$2\tan^{-1}x + \cot^{-1}x = \frac{\pi}{4}.$$

Let θ and ϕ denote the angles $\tan^{-1}x$ and $\cot^{-1}x$ respectively, then taking the tangent of both sides of the equation

$$\tan(2\theta + \phi) = \frac{\tan 2\theta + \tan\phi}{1 - \tan 2\theta \tan\phi} = \frac{\dfrac{2x}{1 - x^2} + \dfrac{1}{x}}{1 - \dfrac{2}{1 - x^2}} = 1.$$

This reduces to

$$x^3 + x^2 + x + 1 = 0.$$

The only real solution is $x = -1$, and this satisfies the original equation if we take $\tan^{-1}x = -\frac{\pi}{4}$ and $\cot^{-1}x = \frac{3\pi}{4}$.

EXERCISES XVIII

1. Solve the following equations:

(a) $2\cos^2 x + 3\sin x = 3.$

(b) $\cos 2x - 3\cos x + 2 = 0.$

(c) $\sin\left(x + \frac{2\pi}{3}\right) + \sin\left(x + \frac{\pi}{3}\right) = \frac{3}{2}.$

(d) $\sin y + \sin 3y = \cos y - \cos 3y.$

(e) $\csc x \cot x = 2\sqrt{3}.$

(f) $\tan^2 x + \cot^2 x = 2.$

2. Solve the following systems of equations.

(a) $\sin y + \cos x = -1,$

$$\frac{\cot^2 y}{\csc^2 y - \csc y} = 2\cos^2 x.$$

(b) $3\sin x + \tan y = \frac{5}{2},$

$$(1 - 2\cot y)\csc x = -2.$$

3. Find x in the following:

(a) $\sin^{-1}x + \cos^{-1} 2x = \frac{\pi}{6};$

(b) $2\sin^{-1}x + \sin^{-1} 3x = \pi;$

(c) $\cos^{-1}x + \tan^{-1}x = \cot^{-1}x.$

4. A body lies on an inclined plane, coefficient of friction $\mu = \frac{1}{\sqrt{3}}$. What should be the inclination of the plane so that the force acting up the plane necessary to move the body is equal to the weight of the body?

5. A ladder leans against a wall, coefficient of friction at both wall and floor is $\mu = \frac{1}{3}$. At what angle will the ladder just stand without slipping?

6. A weight of 100 lb. is supported by two cords, attached at different points on a horizontal beam. If the tensions in the two cords are 60 lb. and 100 lb. find the angles that the cords make with the beam.

Art. 22. Equations Involving Logarithmic Transformations

Equations involving exponential functions are of extreme importance. Since the exponential function is the inverse of the logarithm function, their presence suggests the taking of logarithms. This treatment is always effective if each equation of the system contains only two terms, each term the product of known constants, known powers of unknowns, and known constants with unknown exponents, and if no unknown appears both as base and as exponent.

To illustrate, let it be required to find x when $y = y_1$, from the equation

$$y = ab^{kx},$$

where a, b and k are regarded as known. Taking logarithms,

$$\log y_1 = kx \log b + \log a,$$

$$x = \frac{\log y_1 - \log a}{k \log b}.$$

Here it is immaterial what base is used for the logarithms. The significance of the base was discussed in Article 3 and in Article 12. For practical computation of x the base should be regarded as 10, but frequently in theoretical formulations it is convenient to use the natural base, e.

Again, in the equation above, let it be required to find a and k, given the data $x = 2.7$ when $y = 5.2$ and $x = 3.8$ when $y = 3.4$ and taking $b = e$, the natural base. We have the equations

$$5.2 = a\, e^{2.7k},$$

$$3.4 = a\, e^{3.8k}.$$

Taking logarithms

$$\log 5.2 = 2.7\, k \log e + \log a,$$

$$\log 3.4 = 3.8\, k \log e + \log a.$$

These are linear equations for k and log a. If natural or naperian logarithms are used $\log e = 1$, but the equations are true for any base.

EXERCISES XIX

1. The last term of a geometrical progression is given by the formula $L = a r^{n-1}$, where a is the first term, r the ratio of each term to the preceding, and n the number of terms. Insert four geometrical means between 5 and 23.

2. If T_1 and T_2 are the tensions on a belt as it leaves and approaches a pulley, then in order that the belt shall drive the pulley without slipping of the belt we must have

$$\frac{T_1}{T_2} < e^{a\theta}$$

where a is the coefficient of friction between belt and pulley and θ is the angle (radian measure) at the axis of the pulley subtended by the arc of contact. Find the limit of T_1/T_2 if $a = 0.3$ and $\theta = \pi$. Find the limit of θ if $a = 0.3$ and $T_1 = 2T_2$.

3. Solve the following equations for x, y and z:

$$x^2y^3 = 6z^4, \quad xy^5z^2 = 7, \quad x^5z^3 = y^3.$$

4. Given the formula $pv^n = c$, find n from the data $p = 213$ when $v = 2$ and $p = 23.4$ when $v = 10$.

5. The voltage V at time t across a condenser of capacity K, discharging through a resistance R is given by the equation

$$V = Ae^{-\frac{t}{KR}}$$

where A is a constant. Given $K = \frac{1}{3}\,10^{-6}$, find R from the data $V = 100$ when $t = 0$, $V = 55$ when $t = 1800$.

Art. 23. Approximate Solutions of Equations

In the preceding articles of this chapter we have considered methods for finding solutions of equations and systems of equations of certain types. We have seen that success is attained only when by transformations and combinations of the equations we arrive at a rational algebraic equation in one unknown, the type considered in Art. 19. In Art. 19, we considered a method for finding exactly the rational roots of a polynomial equation, and if the depressed equation is quadratic or binomial, all roots may be found. Now equations and systems of equations frequently arise in which it is not possible to reduce the

problem of solving to that of solving rational algebraic equations, and even when it is possible, the algebraic equations may not yield to the method of Art. 19. In such cases it becomes necessary to adopt some method which yields approximate results, and for practical purposes a graphical treatment possesses great advantages, used either to obtain results directly, or as a guide to other approximate processes.

1. **Graphical Solutions, One Unknown.** Having a single equation in one unknown, we transform the equation so that the function on each side of the equation is as simple as possible, then make graphs of the two functions and observe for what values of the unknown the two functions have the same value, i.e., we note the abscissas of the points where the graphs cross. First making a rough determination, then plotting on a greatly enlarged scale the portions of the graphs in the immediate neighborhood of the intersections, the approximation may be carried to any degree of accuracy required.

In making the graphs, use is made of the principles discussed in Chapter I, and, as tabular and numerical work was there found to be a convenient substitute for graphical work under certain conditions, so will it here. In the actual work of obtaining approximate solutions the use of slide rule, printed tables, tabular computations, and graphs in various combinations will be found advantageous, and the types of equations are so numerous and the adaptations of method so flexible, that no general rules are advisable. But in any case the conception of the intersecting graphs is an effective guide, and in most cases portions of the graph will be found actually useful.

Equations requiring this kind of treatment that arise most frequently are those involving **mixed functions,** that is, involving functions from two or more of the classes of functions mentioned in Chapter I. The method, however, is available for any equation, whether it is impossible to solve by exact methods with proper transformations, or merely inconvenient to solve exactly on account of laborious processes involved, etc.

Suppose it required to solve the equation

$$e^{3x}\sin\left(x - \frac{\pi}{6}\right) = 1.$$

The equation does not yield to any exact method of solution. We may write

$$\sin\left(x - \frac{\pi}{6}\right) = e^{-3x},$$

and making graphs of the two functions we determine the value of x at which they intersect. Looking at tables for sine and exponential functions (for example, Smithsonian Physical Tables, tables 14 and 19), we may observe, even before making graphs, that there is one solution between x = .66 and x = .67. Constructing a brief table, we have,

x	e^{-3x}	$\sin\left(x - \frac{\pi}{6}\right)$
.66	.1381	.1362
.67	.1340	.1459

If we locate these points of the graphs on squared paper (Fig. 22), choosing units as large as the sheet permits, and assume that these

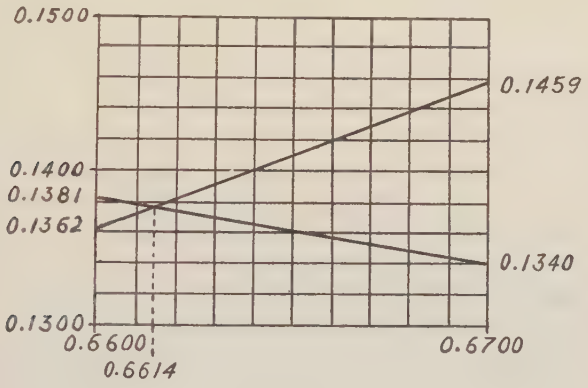


Fig. 22

small portions of the graphs are straight lines, we find that the graphs cross at about x = .6614. This result is accurate to at least three places, sufficient for most practical purposes. Other intersections of the graphs, and therefore other solutions of the equation, are found by making a sketch of the graphs on a small scale, then enlarging the parts near the intersections.

2. **Polynomial Equations.** Equations of this type, with numerical coefficients, yield to the graphical method just considered, but in some cases it may be convenient to adopt a special method based on properties of the polynomial. If the coefficients are rational, it is possible that the equation may have rational roots, in which case the equation may be depressed, giving an equation of lower degree, by

method of Art. 19. Properties of polynomial equations that are useful are stated here, using the notation of Art. 19.

$$P(x) = a_o x^n + a_1 x^{n-1} + \ldots . a_n$$

where the a's are supposed to be given numerically.

(a) If $P(x_1)$ and $P(x_2)$ are of opposite signs there is an odd number of real roots of $P(x) = 0$ between x_1 and x_2.

This is because a polynomial is a continuous function, so that the graph must cross the x-axis an odd number of times between x_1 and x_2.

(b) If $P(x_1)$ and $P(x_2)$ are of the same sign, there can not be an odd number of real roots between x_1 and x_2.

The graph being continuous, if it crosses the x-axis it must cross back again, so that if it crosses at all it must cross an even number of times.

(c) If $a + b\sqrt{-1}$, where a and b are real, is a root of $P(x) = 0$, then $a - b\sqrt{-1}$ is also a root.

If we put $x = a + b\sqrt{-1}$ in $P(x)$, we may separate the real and imaginary terms, so that $P(a + b\sqrt{-1}) = A + B\sqrt{-1}$, where A and B are real. If $a + b\sqrt{-1}$ is a root, then A and B are each zero, since an imaginary number cannot equal a real number. But $P(a - b\sqrt{-1}) = A - B\sqrt{-1}$, since the only change is a different sign for the imaginary expression, $\sqrt{-1}$. Therefore if $a + b\sqrt{-1}$ is a root, its conjugate, $a - b\sqrt{-1}$, is also.

(d) The number of positive real roots for the equation $P(x) = 0$ cannot exceed the number of variations of signs of the coefficients.

By a variation of sign we mean the following of a positive coefficient by a negative one, or the reverse, when the terms are arranged according to the powers of the unknown. The truth of the rule, known as **Descartes' Rule of Signs,** depends on the two facts: First, if an equation has all signs alike it has no positive root, i.e., a sum of positives or a sum of negatives cannot be zero; and second, the polynomial formed by the product $(x - r)\,P(x)$ has at least one more variation of signs than $P(x)$. This last statement is easily verified. If $a_o, a_1, \ldots . a_n$ are the coefficients of $P(x)$, then the coefficients of $(x - r)P(x)$ are $a_o, a_1 - ra_o, a_2 - ra_1, \ldots . , - ra_n$. It is clear that if two successive coefficients, as a_1 and a_2, differ in sign, then a_2 and $a_2 - ra_1$ agree in sign, thus if two terms of $P(x)$ give a variation, the term in the new equation corresponding to the second of these, agrees with it in sign. Also the first coefficient of the new equation is the same as that of the old. Now if a_o and the second term of every pair that produces a variation be written in order, the number of variations is

not altered by omitting the other terms. But these same signs are reproduced in the new equation, with, possibly, other terms between, which may give variations, so that the coefficients down to $a_n - ra_{n-1}$ produce at least as many variations as the original possessed. The additional term, $-ra_n$ produces one more variation, since it certainly differs in sign from the second term involved in the last preceding variation.

To illustrate, take

$$2x^4 - 6x^3 - 3x^2 - 4x + 5 = 0.$$

Here there are two variations, and these variations, appear also in the equation,

$$2x^4 - 6x^3 + 5 = 0,$$

formed by omitting terms which follow terms of like signs. Let us multiply by $x - 1$, writing only coefficients,

2	−6	−3	−4	5	
	−2	6	3	4	−5
2	−8	3	−1	9	−5

The coefficients 2, −8, 9, which correspond to 2, −6, 5, agree with these in sign, and the presence of 3 and −1, actually increases the number of variations, but the number could not be less than two whatever numbers replace 3 and −1. An additional variation is given by −5. The original equation here cannot have more than two real positive roots, hence the new equation cannot have more than three. Applying the rule directly to the new equation we cannot limit the number of real positive roots to less than five.

(e) The number of negative real roots of $P(x) = 0$ cannot exceed the number of variations of signs in $P(-x) = 0$.

For the negative roots of the former equation correspond to positive roots of the latter equation.

The theorems and processes of Article 19 applied in the light of properties (a) − (e) above, should yield results without undue labor. The graph of the function P(x) should be in mind, and usually should be constructed, at least roughly.

Let it be required to find the value of y when $x = \frac{\pi}{4}$ from the equation

$$xy^2 - y = \frac{\sin x}{y^2}.$$

The equation for y is

$$P(y) = .7854\,y^4 - y^3 - .7071 = 0.$$

The coefficients not being rational, there can be no rational roots. There being one variation of signs, the equation can have not more than one positive real root, and in $P(-y)$ the signs are $+\ +\ -$, giving one variation, and one possible negative root. There are, then, two or four imaginary roots. Now by the remainder theorem, using synthetic division, we find $P(0) = -.707$, $P(1) = -.922$ and $P(2) = 3.859$. There is a real root between 1 and 2, and the location of these points of the graph (Fig. 23) seems to indicate that the root may

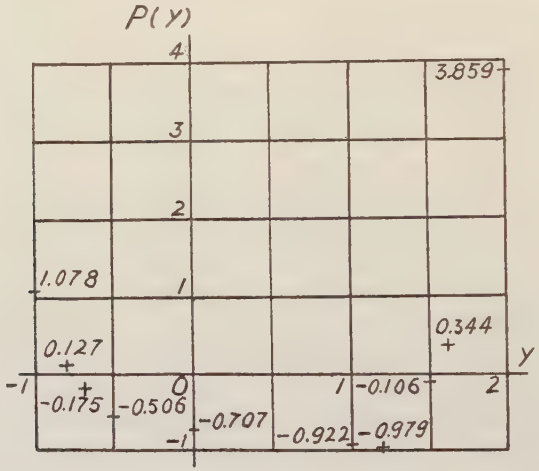


Fig. 23

be between 1.2 and 1.3. But we find $P(1.2) = -.979$, and proceeding further, $P(1.5) = -.106$ and $P(1.6) = .344$. The root is between 1.5 and 1.6, and using the ordinary method of interpolation we have for this root,

$$y = 1.5 + \frac{106}{106 + 344} \cdot \frac{1}{10} = 1.524.$$

Now turning to $P(-y)$ and trying $y = 1$, we find $P(-1) = 1.078$, indicating that the negative root is between -1 and 0, and we compute $P(-.5) = -.506$, $P(-.7) = -.175$ and $P(-.8) = .127$, then by ordinary interpolation,

$$y = -.7 - \frac{175}{175 + 127} \cdot \frac{1}{10} = -.758$$

The last figure in each root is probably wrong. An alternative method for solving this equation would be to write

$$.7854y^4 = y^3 + .7071,$$

and make graphs of the two equated functions, using slide rule, tables of powers, etc., then determine the points of intersection.

3. **Simultaneous Equations.** In Article 20, we saw that if a system of equations can be reduced to a system of rational algebraic equations it is possible to eliminate all but one unknown, and obtain a polynomial equation in this unknown. This may be done for all the

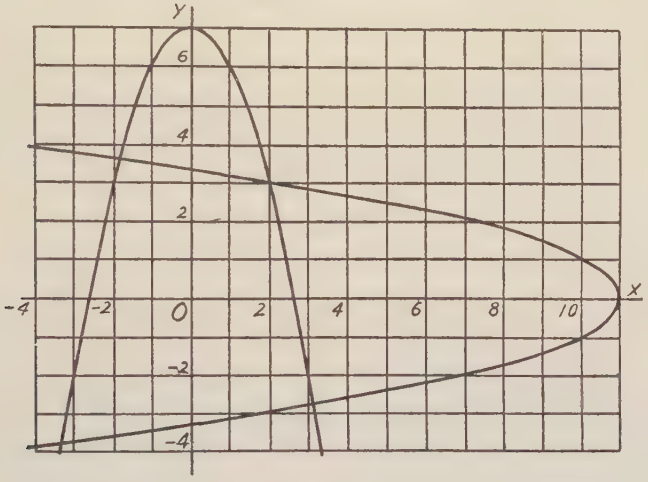


Fig. 24

unknowns in turn. But systems arise which cannot be made rational algebraic equations, and which yield to no method of elimination; and even if elimination is possible, the resulting equations in the single unknowns may require approximate methods of solving. In the case of two equations in two unknowns a graphical method is available without elimination.

Given two equations connecting the two unknowns, x and y, each equation determines y as a function of x, and we may make graphs of the two functions, and determine the values of x and y where the graphs cross.

Consider the equations $x^2 + y = 7$, $x + y^2 = 11$. Here elimination

of either y or x leads to a fourth degree equation in the other unknown, the equations being,

$$x^4 - 14x^2 + x + 38 = 0,$$

$$y^4 - 22y^2 + y + 110 = 0.$$

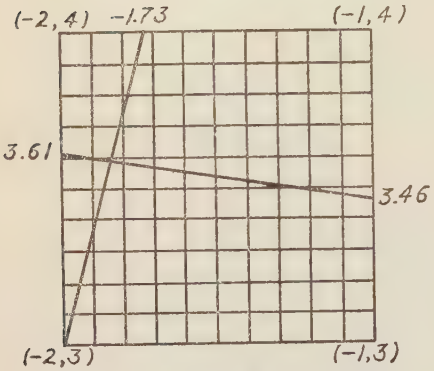


Fig. 25

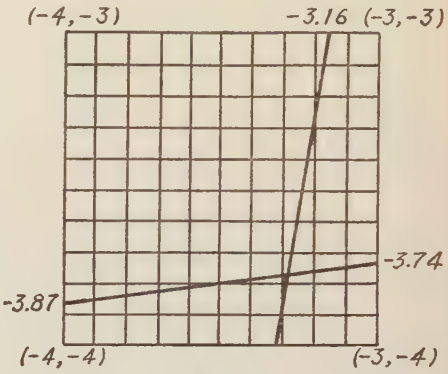


Fig. 26

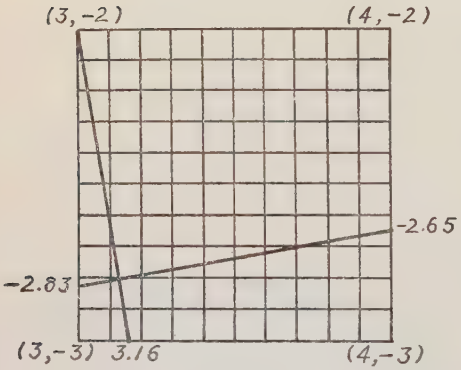


Fig. 27

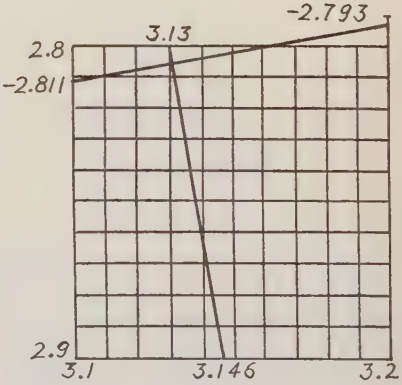


Fig. 28

Each of these has one rational root, x = 2, y = 3, and a reasonable plan is to solve for the remaining values of x by the method in 2, above, then obtain y by the equation $y = 7 - x^2$.

We may, however, make graphs from the two original equations, as shown in Fig. 24, locating the four points of intersection. One of these is the point where x = 2, y = 3, and by drawing on a large scale

the unit squares in which the other intersections occur (Figures 25, 26, 27), we estimate the remaining three solutions to be $x = -1.84$, $y = 3.60$; $x = -3.29$, $y = -3.78$; and $x = 3.13$, $y = -2.80$. If greater accuracy is desired, we may plot on a still larger scale the small regions in which these intersections are found to occur. The construction is shown for the fourth quadrant solution (Fig. 28), leading to the estimated values $x = 3.131$, $y = -2.805$.

This method is available for any two equations in two unknowns. If elimination is possible, the choice of methods will be governed by the particular conditions of the problem, but if elimination is impossible or inexpedient, the graphical method furnishes a practical means of securing approximate solutions.

EXERCISES XX

1. Solve each of the following equations:

(a) $x \sin x - \cos x = 0$, (least positive root);
(b) $(1 - x) \log_{10} x = -1$;
(c) $e^{2x}(3x + 4x^3) = 5$.

2. Find the real roots of the following equations:

(a) $x^3 - 5x^2 + 2x + 7 = 0$;
(b) $x^4 + x^3 - x^2 - 2x - 2 = 0$;
(c) $x^3 - 3x - \frac{33}{4} = 0$.

3. Find the real solutions of the following systems of equations:

(a) $y^2 = 8x^3$, $x^2 + y = 11$; (b) $y = \sin^{-1}\frac{x}{y}$, $x = ye^{4x}$; (c) $\sin x = \tan y$, $x + y = \frac{\pi}{2}$.

4. The expression for current, i, in an electrical circuit as a function of the time t, is

$$i = \frac{E}{R} + Ce^{-\frac{R}{L}t},$$

where E is the impressed electro-motive force, R the resistance of the circuit, L is the inductance, and C is a constant depending on the initial conditions. As t increases the last term diminishes, and the current approaches the limiting value $\frac{E}{R}$. If $i = 10$ when $t = 0$, what

is the value of C? If C has this value, and E = 100 and L = 0.01, find R so that when t = 0.1, i is 98% of its limiting value.

5. If a chain of length s feet is suspended between points at the same level and distant d feet apart, we have

$$s = 2\frac{H}{w}\sinh\frac{wd}{2H} = \frac{H}{w}\left(e^{\frac{wd}{2H}} - e^{-\frac{wd}{2H}}\right).$$

Find H if s = 200, d = 150, and w = 10. Here w is weight per foot of chain and H is the horizontal component of tension.

BOOKS FOR SUPPLEMENTARY READING

First Course in the Theory of Equations, L. E. Dickson (Wiley, 1922).
Elementary Theory of Equations, L. E. Dickson (Wiley, 1914).
Theory of Equations, Burnside and Panton (Longmans, 1899).
Introduction to Higher Algebra, Maxime Bocher (Macmillan, 1919), especially Chapters I–IV.
Graphical and Mechanical Computation, Joseph Lipka (Wiley, 1918).
The Calculus of Observations, Whitaker and Robinson (Blackie and Sons, 1924), Chapters V and VI.

CHAPTER VI

THE PROCESS OF INTEGRATION

Art. 24. Integration by Substitution

In Chapter IV we gave attention to the meaning of integration, while in the present chapter we consider the methods or processes by which integration is accomplished. The equation

$$\int \phi(x)dx = f(x)$$

was explained in Art. 15 to mean that an operation, denoted by the symbol $\int \ldots dx$, is applied to $\phi(x)$ to obtain $f(x)$, and that this operation is the inverse of finding the derivative. The two sentences, "the integral of $\phi(x)$ is $f(x)$," and "the function whose derivative is $\phi(x)$ is $f(x)$" are accurate interpretations of the equation. Now in Art. 18 we gave a meaning to the notation $\phi(x)dx$, and called it the differential of $f(x)$, and from this point of view we may regard the equation as indicating that an operation, denoted by $\int$, is applied to the differential $\phi(x)dx$ to obtain $f(x)$. The language in this case is, "the integral of $\phi(x)dx$ is $f(x)$," or "the function whose differential is $\phi(x)dx$ is $f(x)$." It is at once apparent that the word integral is given different meanings in the two cases, but that the two interpretations of the equation are entirely consistent. In the present chapter, where the process of integration is under consideration, it is advantageous to regard this process as the inverse of finding the differential.

Given an expression,

$$\int \phi(x)dx,$$

and desiring to perform the indicated operation, we attempt to identify the expression with one of the standard forms in Art. 16. If this cannot be done directly, we may make a substitution, putting a new variable, say u, for some function of x, say $u = \theta(x)$, and from this we have

$du = \theta'(x)dx$, so that we are able to solve for x and dx in terms of u and du, and obtain for the integral a new expression in the form

$$\int F(u)du.$$

If the function $u = \theta(x)$ is skilfully chosen, the new expression may be more easily identified with a standard form. This method is equivalent to the **method of substitution** explained in Art. 16, and the examples given there serve to illustrate its operation.

In the case of a definite integral we have two alternatives. Having made the substitution we may integrate and obtain a function of u, then transform back in terms of x before evaluating at the limits of integration; or we may provide new limits of integration, the values that u takes when x takes the values of the old limits, and evaluate the integral in terms of u at these limits.

For example, let us evaluate the integral

$$\int_0^{\frac{1}{2}} \sqrt{1 - 4x^2}dx.$$

We substitute $2x = \sin u$, then $dx = \frac{1}{2}\cos u\, du$, and the indefinite integral involved reduces to

$$\frac{1}{2}\int \sqrt{1 - \sin^2 u}\cos u\, du = \frac{1}{2}\int \cos^2 u\, du,$$

and by trigonometric reduction this is

$$\frac{1}{4}\int (1 + \cos 2u)\, du = \frac{1}{4}\int du + \frac{1}{8}\int \cos 2u\, d(2u).$$

Integrating, this becomes

$$\frac{1}{4}u + \frac{1}{8}\sin 2u = \frac{1}{4}u + \frac{1}{4}\sin u \cos u.$$

In terms of x this is

$$\frac{1}{4}\sin^{-1}2x + \frac{x}{2}\sqrt{1 - 4x^2},$$

and evaluating at the limits of integration we have the value $\frac{\pi}{8}$.

But by the alternative method we observe that $u = 0$ when $x = 0$, and $u = \frac{\pi}{2}$ when $x = \frac{1}{2}$, and writing

$$\frac{1}{2}\int_0^{\frac{\pi}{2}} \cos^2 u \, du,$$

and integrating as before, we have

$$\left[\frac{1}{4}u + \frac{1}{8}\sin 2u\right]_0^{\frac{\pi}{2}} = \frac{\pi}{8}.$$

Success in the use of the method of substitution depends upon the skill of the person using it, and this skill is obtained only by practice. Often the method is used in combination with other methods explained in the following articles. It is not practicable to list the types of functions that are best handled by this method, but attention is directed to a few of the more common and obvious cases where substitution is suggested in the following exercises.

EXERCISES XXI

1. Find the function denoted by

$$\int (4 - 3x)^{\frac{3}{2}} dx.$$

(Substitute $4 - 3x = u$.)

2. Integrate the following:

(a) $\int \sqrt{4x - 2}\, dx$; (b) $\int \frac{dx}{\sqrt{5 - 7x}}$; (c) $\int (2x^2 - 5)^3 x dx$.

3. Perform the integration

$$\int \frac{dv}{\sqrt{4 - v^2}}.$$

(Substitute $\frac{v}{2} = u$ and use form 12 of Art. 16.)

4. Integrate the following:

(a) $\int \frac{dx}{1 + 4x^2}$; (b) $\int \frac{dx}{x\sqrt{x^2 - 9}}$; (c) $\int \frac{dx}{\sqrt{4x^2 + 9}}$; (d) $\int \frac{dx}{1 - 9x^2}$.

5. Evaluate $\int_3^5 \frac{dx}{2x - 5}$. (Substitute $2x - 5 = u$.)

6. Evaluate (a) $\int_8^{10} \frac{dx}{x - 7}$; (b) $\int_{-4}^{0} \frac{dx}{1 - 3x}$; (c) $\int_0^2 \frac{x dx}{1 + x^2}$.

7. Integrate $\int \sqrt{1 - 9x^2}\, dx$. (Substitute $3x = \sin u$.)

8. Integrate $\int \sqrt{4 - x^2}\, dx$.

9. Integrate (a) $\int x \cos x^2 dx$; (b) $\int x \sqrt{1 + 4x^2} dx$;

(c) $\int x \csc^2(1 - x^2) dx$. (In the first substitute $x^2 = u$.)

10. Integrate (a) $\int xe^{1-x^2} dx$; (b) $\int x^2 e^{4x^3-2} dx$. (In the former substitute $1 - x^2 = u$.)

11. Integrate (a) $\int e^{\sin x} \cos x\, dx$; (b) $\int \cos^3 x \sin x\, dx$.

Art. 25. Integration by Parts

By Rule II for finding derivatives we have

$$\frac{d}{dx}(uv) = v\frac{du}{dx} + u\frac{dv}{dx},$$

where u and v are any functions of x. From this rule we have, obviously,

$$uv = \int vdu + \int udv,$$

and when we transpose to the form

$$\int udv = uv - \int vdu$$

we have a plan for integration of products which is known as **integration by parts.**

Given a function to integrate,

$$\int f(x)dx,$$

the plan is to separate f(x) into two factors, one of which can be recognized as the derivative of some function, v, then denoting the other factor by u, we arrive at the value

$$uv - \int vdu,$$

where the new integration may be easily performed, either by another application of the plan of integration by parts, or by some other method.

The plan seems similar to the method of substitution as explained in Article 16, but is essentially different in that there the one factor must be a function of the quantity of which the other is a derivative, as if u here were a function of v, while in integration by parts u and v are independent functions.

Attention should be called to the fact that, by the formula above,

$$\int vdu = uv - \int udv,$$

so that,

$$\int udv = uv - \int vdu = uv - (uv - \int udv).$$

This is seen to be an identity, and furnishes no information concerning the value of the integral. In repeated application of the plan the factors u and dv must be so chosen that this kind of identity does not occur.

Again, in repeated application of the plan of integration by parts, if the integrals that arise at successive stages of the work grow more involved and difficult, a new start should be made, with a different choice of factors. In some cases a second application of the plan may lead to a new integral exactly equal to the original except for a constant factor. In this case the relation is not an identity, and the equation may be solved for the value of the integral without further integration.

For example, take

$$\int e^x \sin x \, dx.$$

Put $u = e^x$, $dv = \sin x \, dx$, then $v = -\cos x$,

$$\int e^x \sin x \, dx = -e^x \cos x + \int \cos x \, e^x dx.$$

Now take $u = e^x$, $dv = \cos x \, dx$, then $v = \sin x$,

$$\int \cos x \, e^x dx = e^x \sin x - \int \sin x \, e^x dx,$$

and we have, combining the two equations, the desired result,

$$2 \int e^x \sin x \, dx = e^x(\sin x - \cos x) + C.$$

If at the second stage of the work we had taken $u = \cos x$, $dv = e^x dx$, then $v = e^x$, we should have had

$$\int \cos x\, e^x dx = e^x \cos x + \int e^x \sin x\, dx,$$

which is identical with the result of the first stage, and leads to no useful result.

Again consider the integral

$$\int \log x\, dx.$$

This does not occur among the standard forms, but putting $u = \log x$, $dv = dx$, we have

$$\int \log x\, dx = x \log x - \int x \frac{dx}{x} = x \log x - x + C.$$

EXERCISES XXII

1. Integrate: (a) $\int \log(5x - 8)dx$; (b) $\int e^{2x} \cos x\, dx$;

(c) $\int e^x \sin 3x\, dx$; (d) $\int e^{-2x} \cos 3x\, dx$.

2. Evaluate: (a) $\int_0^{\pi} x \sin x\, dx$; (b) $\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} x \cos 2x\, dx$;

(c) $\int_0^1 x\, e^x dx$; (d) $\int_{-1}^{0} x^2 e^{3x} dx$; (e) $\int_1^2 x\, e^{-2x} dx$.

3. Integrate: (a) $\int x \log x\, dx$; (b) $\int x \cosh x\, dx$. (See Exercises IX, 4, Article 12.)

4. Integrate $\int \sin^{-1} x\, dx$. (First apply the method of substitution, putting $u = \sin^{-1} x$, then integrate the result by parts.)

5. Integrate: (a) $\int \cos^{-1} x\, dx$; (b) $\int \sinh^{-1} x\, dx$

Art. 26. Use of Trial Integrals

A method of obtaining integrals that is very useful in practice is that of writing a function which is known to have the form of the integral, but which contains certain unknown coefficients which are to be de-

termined. This function, known as a **trial integral**, must have a derivative identical with the given integrand, and this condition is used to determine the values of the unknown coefficients. The method is successful in case the given integrand is such that we may, without actually integrating, arrive at a correct form of the integral, except for the values of certain coefficients. Fortunately this is possible in the case of many of the functions that occur most frequently in practice.

If a function is such that when its successive derivatives are formed a certain set of terms are obtained so that any derivative is the sum of a group of these terms, each multiplied by a constant, then, of course integration, as well as differentiation, will introduce no new terms, but merely new coefficients. The trial integral may then be taken as the sum of all distinct terms that occur in the integrand and its successive derivatives, each term having an arbitrary coefficient.

For example take

$$\int x^2e^x dx.$$

The integrand and its successive derivaties are: x^2e^x; $x^2e^x + 2xe^x$; $x^2e^x + 4xe^x + 2e^x$; $x^2e^x + 6xe^x + 6e^x$. Further differentiation will introduce no new terms, but merely new coefficients. Here, then, is a group of terms, x^2e^x, xe^x, e^x, such that differentiation only leads to repetition of the terms with different coefficients, therefore, integration of any of these terms will lead to terms of the same group, with constant coefficients. For a trial integral take

$$ax^2e^x + bxe^x + ce^x$$

where a, b and c are to be determined. Equating the derivative of this to the given integrand we have

$$ax^2e^x + (2a + b)xe^x + (b + c)e^x = x^2e^x.$$

We see that $a = 1$, $2a + b = 0$, $b + c = 0$, whence $b = -2$, $c = 2$. The integral is, therefore,

$$\int x^2e^x dx = x^2e^x - 2xe^x + 2e^x + k.$$

Any exponential function ae^{kx}, when differentiated repeatedly gives terms of the same kind, differing only in the coefficients. Any function of the form of a sin kx or a cos kx gives only terms of one of these forms when differentiated repeatedly. Repeated differentiation of ax^n, where

n is a positive integer, leads to lower powers of x, and finally to zero. From these considerations, it is not difficult to see that any integrand consisting of a sum of products of exponentials, sines, cosines, and polynomials may be treated by the use of a trial integral.

EXERCISES XXIII

1. Integrate the following:

(a) $\int x \sin 2x \, dx$; (b) $\int x \cos 4x \, dx$; (c) $\int (1 - x)\sin \frac{x}{2} \, dx$.

2. Find the following integrals:

(a) $\int e^x \sin \frac{x}{3} \, dx$; (b) $\int e^{-3x}(\cos x - 2 \sin x)dx$.

3. Evaluate the following:

(a) $\int_0^1 e^x(x^2 - 3x)dx$; (b) $\int_1^3 e^{-\frac{x}{2}}(x^3 - 4x + 2)dx$.

4. Integrate:

(a) $\int x \sinh x \, dx$; (b) $\int \sin x \cosh x \, dx$.

Art. 27. Integration of Rational Fractions

The integration of polynomials presents no difficulties, but the treatment of rational fractions, the ratios of polynomials, is a matter that requires special attention. If a rational fraction has for numerator a polynomial of degree higher than, or equal to, that of the denominator, then by dividing the numerator by the denominator the function is reduced to a mixed quantity, the sum of a polynomial and a rational fraction whose numerator is of lower degree than the denominator. For example, given the rational fraction

$$\frac{x^4 + 2x^3 - 3x + 2}{3x^3 - 6x + 2},$$

we find, by dividing, the quotient $\frac{1}{3}(x + 2)$ and the remainder $2x^2 + \frac{1}{3}x + \frac{2}{3}$, so that the fraction is equal to

$$\frac{1}{3}(x + 2) + \frac{2x^2 + \frac{1}{3}x + \frac{2}{3}}{3x^3 - 6x + 2}.$$

It will be sufficient then, to treat the problem of finding the integral of a rational fraction whose numerator is of lower degree than the denominator. Suppose it required to integrate such a function, and suppose the denominator to be of degree n. By the principles concerning polynomial equations, Art. 19, there exist n linear factors of this polynomial, and the denominator may be written in the form

$$a_0(x - r_1)(x - r_2)(x - r_3) \ . \ . \ . \ . (x - r_n),$$

where $r_1, r_2 \ . \ . \ . \ r_n$ are constants, and a_0 may be divided out. Consider now the sum of the **partial fractions,**

$$\frac{C_1}{x - r_1} + \frac{C_2}{x - r_2} + \ . \ . \ . \ . \frac{C_n}{x - r_n}.$$

If these denominators, the factors of the original denominator, are **all distinct,** then the sum of these partial fractions will have for denominator the denominator of the original fraction, and for numerator it will have a polynomial of degree $n - 1$, the coefficients of which are linear functions of the C's. Equating these coefficients to the coefficients of like powers of x in the numerator of the original fraction we obtain n equations from which we may find values of the C's which make the sum of partial fractions equal to the original fraction. As illustration we may take

$$\frac{\frac{1}{2}(17 - x)}{(x - 2)(x + 1)(x + 3)} = \frac{C_1}{x - 2} + \frac{C_2}{x + 1} + \frac{C_3}{x + 3}.$$

In adding the three fractions on the right we find for the numerator

$$C_1(x + 1)(x + 3) + C_2(x - 2)(x + 3) + C_3(x - 2)(x + 1),$$

which reduces to

$$(C_1 + C_2 + C_3)x^2 + (4C_1 + C_2 - C_3)x + (3C_1 - 6C_2 - 2C_3)$$

and if this is identical with $\frac{1}{2}(17 - x)$ we must have

$$\begin{aligned} C_1 + \ C_2 + \ C_3 &= 0, \\ 4C_1 + \ C_2 - C_3 &= -\frac{1}{2} \\ 3C_1 - 6C_2 - 2C_3 &= \frac{17}{2} \end{aligned}$$

This gives $C_1 = \frac{1}{2}$, $C_2 = -\frac{3}{2}$, $C_3 = 1$, and the integral of the original fraction is

$$\frac{1}{2}\int \frac{(17 - x)dx}{(x-2)(x+1)(x+3)} = \frac{1}{2}\int \frac{dx}{x-2} - \frac{3}{2}\int \frac{dx}{x+1} + \int \frac{dx}{x+3} =$$

$$\frac{1}{2}\log(x-2) - \frac{3}{2}\log(x+1) + \log(x+3) = \log\sqrt{\frac{(x-2)(x+3)^2}{(x+1)^3}}.$$

If the factors of the denominator are **not all distinct,** then the common denominator of the fractions in the sum above is not the denominator of the original fraction, showing that the sum as written is not a suitable form in which to express the given function. A simple modification of the process is, however, effective. If several factors are equal, that is, a power of one factor occurs, say $(x - r)^m$, where m is less than or equal to n, then the corresponding m terms in the sum as previously written are replaced by the terms

$$\frac{a_1}{x-r} + \frac{a_2}{(x-r)^2} + \cdots \frac{a_m}{(x-r)^m}.$$

The common denominator of the fractions on the right will now be the denominator of the given fraction, and the n constants are available to fulfil the conditions that make the sum have the required value.

To illustrate this point, take

$$\frac{x^3 + 2x^2 + 5x}{(x-1)(x+1)^3} = \frac{C_1}{x-1} + \frac{C_2}{x+1} + \frac{C_3}{(x+1)^2} + \frac{C_4}{(x+1)^3}.$$

The numerator, on adding the fractions on the right, is

$$C_1(x+1)^3 + C_2(x-1)(x+1)^2 + C_3(x-1)(x+1) + C_4(x-1)$$

or collecting powers of x,

$$(C_1 + C_2)x^3 + (3C_1 + C_2 + C_3)x^2 + (3C_1 - C_2 + C_4)x + (C_1 - C_2 - C_3 - C_4)$$

and this is $x^3 + 2x^2 + 5x$ if

$$\begin{aligned} C_1 + C_2 &= 1, \\ 3C_1 + C_2 + C_3 &= 2, \\ 3C_1 - C_2 + C_4 &= 5, \\ C_1 - C_2 - C_3 - C_4 &= 0. \end{aligned}$$

From this we have $C_1 = 1$, $C_2 = 0$, $C_3 = -1$, $C_4 = 2$. The resulting fractions are readily integrated since any integral of the form

$$\int \frac{c\,dx}{(ax + b)^n}$$

may be found by use of the substitution $u = ax + b$.

The method as indicated is applicable in all cases, providing the linear factors can be found. If, however, **imaginary factors** are present another modification of the method is advantageous, based on the fact that

$$\frac{a_1}{x - r_1} + \frac{a_2}{x - r_2} = \frac{(a_1 + a_2)x - (a_1 r_2 + a_2 r_1)}{x^2 - (r_1 + r_2)x + r_1 r_2},$$

that is, the sum of any two simple fractions of the kind considered is a fraction of the form

$$\frac{C_1 x + C_2}{x^2 + px + q}.$$

Imaginary factors always enter in pairs, the product of the two being a real quadratic, as the denominator just written, and the two terms of the sum which corresponds to the two imaginary factors may be replaced by a fraction of the form just written. If such factors are repeated, the required form is obtained by writing as denominators the powers of the quadratic up to the power which occurs in the given denominator, each with a numerator of the form $C_1 x + C_2$.

As a final illustration let it be required to integrate

$$\int \frac{15x^2 + 8x - 16}{x^3(x^2 + 3x + 4)^2}\,dx.$$

We write

$$\frac{15x^2 + 8x - 16}{x^3(x^2 + 3x + 4)^2} = \frac{C_1}{x} + \frac{C_2}{x^2} + \frac{C_3}{x^3} + \frac{C_4 x + C_5}{x^2 + 3x + 4} + \frac{C_6 x + C_7}{(x^2 + 3x + 4)^2},$$

and the values of the C's may be found as before.

Here we meet the problem of integrating an expression of the form

$$\int \frac{(ax + b)dx}{(x^2 + px + q)^n}.$$

If we substitute $u = x + \frac{p}{2}$, this may be written

$$a\int \frac{udu}{\left[u^2 + \left(q - \frac{p^2}{4}\right)\right]^n} + \left(b - \frac{ap}{2}\right)\int \frac{du}{\left[u^2 + \left(q - \frac{p^2}{4}\right)\right]^n}$$

The first of these is integrated at once by the substitution $v = u^2 + \left(q - \frac{p^2}{4}\right)$. The treatment of the second integral is based on a **reduction formula** which may be derived as follows:

$$\frac{d}{dx}x(x^2 + k)^m = (x^2 + k)^m + 2mx^2(x^2 + k)^{m-1}$$

and adding and subtracting $2mk(x^2 + k)^{m-1}$ on the right, this becomes

$$(x^2 + k)^m + 2m(x^2 + k)(x^2 + k)^{m-1} - 2mk(x^2 + k)^{m-1}$$
$$= (2m + 1)(x^2 + k)^m - 2mk(x^2 + k)^{m-1}.$$

We may write, therefore,

$$x\,(x^2 + k)^m = (2m + 1)\int (x^2 + k)^m dx - 2mk\int (x^2 + k)^{m-1}dx,$$

or, solving for the integral last written, and putting $m = 1 - n$,

$$\int \frac{dx}{(x^2 + k)^n} = \frac{1}{2(n - 1)k} \cdot \frac{x}{(x^2 + k)^{n-1}} - \frac{3 - 2n}{2(n - 1)k}\int \frac{dx}{(x^2 + k)^{n-1}}.$$

To obtain the value of the integral on the left we apply this reduction formula repeatedly, each time reducing by unity the exponent of the denominator of the integrand, finally arriving at the integral

$$\int \frac{dx}{x^2 + k} = \frac{1}{\sqrt{k}}\tan^{-1}\frac{x}{\sqrt{k}} + C_1 = \frac{1}{2\sqrt{-k}}\log\frac{x + \sqrt{-k}}{x - \sqrt{-k}} + C_2$$
$$= \frac{1}{\sqrt{-k}}\tanh^{-1}\frac{x}{\sqrt{-k}} + C_3.$$

The integration here is performed by use of forms 13 and 20 of Art. 15, the former being useful if k is positive, the latter if k is negative. The reduction formula is seen to apply to the integral in hand, by putting $u = x$ and $q - \frac{p^2}{4} = k$.

We have found, then that any rational algebraic function of one

variable can be integrated if, when written in the form of a fraction, the denominator can be factored into its linear and quadratic factors.

EXERCISES XXIV

1. The following integrations are to be verified by form 20, Art.16, and also by use of partial fractions and form 17 of Art. 16. In both methods a substitution is required.

(a) $\int \frac{dx}{x^2 - 9} = \frac{1}{6} \log \frac{x-3}{x+3} + C$; (b) $\int \frac{dx}{x^2 + 2x - 3} = \frac{1}{4} \log \frac{x-1}{x+3} + C$;

(c) $\int \frac{dx}{4x^2 + 4x - 8} = \frac{1}{12} \log \frac{x-1}{x+2} + C$.

2. Obtain the following integrals by use of a substitution and form 17, Art. 16, also by partial fractions:

(a) $\int \frac{x+1}{x^2 + 2x - 3} dx = \log \sqrt{x^2 + 2x - 3} + C$.

(b) $\int \frac{xdx}{4x^2 - 25} = \frac{1}{8} \log (4x^2 - 25) + C$.

3. Verify the following:

(a) $\int \frac{x-1}{x^2 + 2x + 2} dx = \log \sqrt{x^2 + 2x + 2} - 2 \tan^{-1}(1 + x) + C$;

(b) $\int \frac{3x - 6}{x^2 - 5x + 6} dx = \log(x - 2) + 2 \log(x - 3) + C$;

(c) $\int \frac{x^2 + 2}{x^2 + 1} dx = x + \tan^{-1}x + C$.

4. Apply partial fractions to obtain the integrals indicated.

(a) $\int \frac{x^2 + x - 1}{x^3 + x^2 - 6x} dx$; (c) $\int \frac{(3x - 2)dx}{x(x + 1)^3}$;

(b) $\int \frac{(t^2 + pq)dt}{t(t - p)(t + q)}$; (d) $\int \frac{(2x^2 - 3x - 3)dx}{(x - 1)(x^2 - 2x + 5)}$.

ART. 28. LIMITATIONS OF THE PROCESS OF INTEGRATION

1. **Integrand continuous.** If $\phi(x)$ is continuous from $x = a$ to $x = b$, then there exists a continuous function, say $f(x)$, on the range from a to b, such that

$$\int \phi(x)dx = f(x).$$

To prove this, consider a graph (Fig. 29) of the function $\phi(x)$. We shall define the function f(x) by saying what its value shall be for any x on the required range. Let f(c), the value at c, be any constant. Let $f(x) = f(c) + A_x$, where A_x is the area under the curve and between c and x. The function is continuous, since $A_{(x+\Delta x)} = A_x + \Delta A$ approaches A as Δx approaches zero; and it is an integral of $\phi(x)$, since

$$\frac{d}{dx}f(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{A_{(x+\Delta x)} - A_x}{\Delta x} = \phi(x).$$

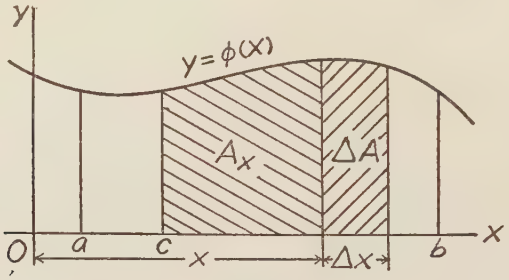


Fig. 29

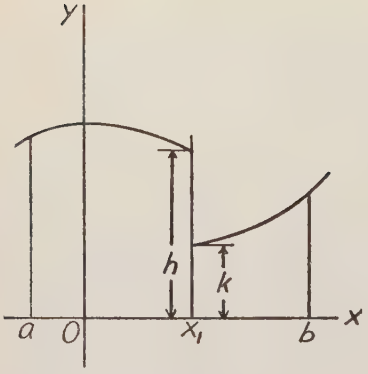


Fig. 30

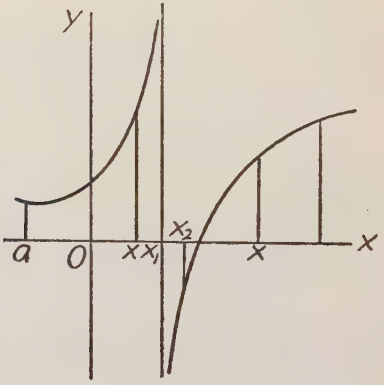


Fig. 31

The function f(x) so defined contains the arbitrary constant C = f(c), which is the constant of integration.

2. **Integrand finitely discontinuous.** Suppose the function $\phi(x)$ to be continuous from a to b (Fig. 30) except at one point, x_1, and suppose that as x approaches x_1 from the left $\phi(x)$ approaches the value h, while as x approaches x_1 from the right $\phi(x)$ approaches k, where h and k are different. Here also we have a continuous function which is the

integral of the given function on the range from a to b. Let $f(a) = C$, an arbitrary constant, and let $f(x) = C + A_x$, where A_x is the area under the curve between ordinates at a and x. The function so defined is continuous, and its derivative is $\phi(x)$ for every x except at x_1. At x_1 the derivative of $f(x)$ does not exist, but the fraction

$$\frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x}$$

approaches h as Δx approaches zero from the negative side, while it approaches k as Δx approaches zero from the positive side.

It is clear that a function $\phi(x)$ may have finite discontinuities at any finite number of points on its range and still have a continuous integral.

3. **Integrand with vertical asymptote.** Suppose $\phi(x)$ to be continuous at every point from a to b except at a point x_1, and suppose that it has a vertical asymptote at x_1 (Fig. 31). Now we may take the integral

$$\int_a^x \phi(x)dx = f(x) \tag{1}$$

where x is between a and x_1, then, of course, the value $f(x)$ is the area under the curve between a and x. As x approaches x_1, this area may approach a definite limit, or it may increase indefinitely, depending on the character of the function $\phi(x)$. If this limit is definite, we may take

$$\lim_{x \to x_1} \int_a^x \phi(x)dx = f(x_1), \tag{2}$$

thus assigning a definite value to $f(x)$ at x_1. Again the integral

$$\int_{x_2}^x \phi(x)dx;$$

where x_2 and x are between x_1 and b, may approach a definite limit or may increase or diminish indefinitely as x_2 approaches x_1. If this limit and the previous one are definite we may take, for any x between x_1 and b,

$$f(x) = f(x_1) + \lim_{x_2 \to x_1} \int_{x_2}^x \phi(x)dx. \tag{3}$$

The function $f(x)$ is assigned definite values on the interval from a to b by equations (1), (2) and (3), and is continuous on this interval. If either of the limits referred to does not exist, i.e., if either quantity concerned increases or diminishes indefinitely, the integral still exists, and is

continuous except at x_1, but at this point it has a vertical asymptote.

The vital importance of this lies in the fact that, **when the integral is discontinuous, integration between limits that lie on opposite sides of the point of discontinuity becomes impossible,** and failure to observe the discontinuity may lead to most absurd results. On the other hand, **when the integral function is continuous, evaluation between limits of integration may be applied with assurance, even though the integrand is discontinuous.**

Example 1. Take for integrand the function $y = \frac{1}{(x-1)^{\frac{2}{3}}}$. This has a vertical asymptote at $x = 1$. The integral is

$$\int \frac{dx}{(x-1)^{\frac{2}{3}}} = 3\,(x-1)^{\frac{1}{3}} + C,$$

This function is continuous at $x = 1$, and the definite integral

$$\int_0^2 \frac{dx}{(x-1)^{\frac{2}{3}}} = \left[3(x-1)^{\frac{1}{3}}\right]_0^2 = 6,$$

is a correct value for the area under the graph of the integral between 0 and 2.

Example 2: Take for integrand $y = \frac{1}{(x-1)^2}$, which has a vertical asymptote at $x = 1$. We have for the integral function

$$\int \frac{dx}{(x-1)^2} = -\frac{1}{x-1} + C,$$

which also has an asymptote at $x = 1$. A definite integral process applied as usual,

$$\int_0^2 \frac{dx}{(x-1)^2} = \left[-\frac{1}{x-1}\right]_0^2 = -2,$$

is here absurd, for

$$\int_0^x \frac{dx}{(x-1)^2} = -\frac{1}{x-1} - 1 = \frac{-x}{x-1},$$

and this is correct if x is less than 1, but as x approaches 1, this increases indefinitely. Also,

$$\int_x^2 \frac{dx}{(x-1)^2} = -1 + \frac{1}{x-1} = \frac{2-x}{x-1},$$

is correct if x is greater than 1, but as x approaches 1, this increases indefinitely. Furthermore, the area from 0 to 2 all lies above the x-axis, yet the value found above is negative, a result obviously fallacious.

4. **Infinite limits of integration.** We have just seen that the area bounded by a curve, the x-axis, a vertical asymptote and an ordinate to the curve may be finite. If the variables are interchanged, the asymptote becomes horizontal, and the expression for the area becomes an integral with an infinite limit of integration. If the function $\phi(x)$ has the x-axis as an asymptote, so that the function approaches zero as x increases indefinitely, then the area under the curve (Fig. 32) from a to x is

$$\int_a^x \phi(x)dx,$$

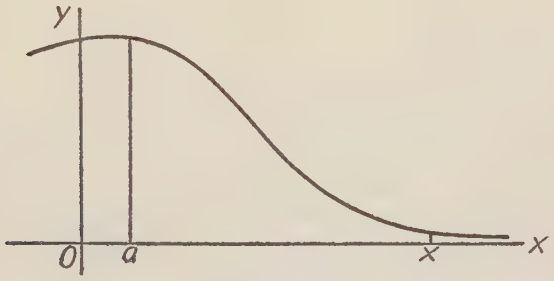


FIG. 32

and if this approaches a definite limit as x increases indefinitely, this limit is the total area under the curve to the right of a. We may write then,

$$\int_a^\infty \phi(x)dx = \lim_{x\to\infty} \int_a^x \phi(x)dx,$$

But if the integral does not approach a definite limit the expression on the left has no real meaning, and the situation is stated in the words "the integral does not exist," or "the integral is not finite."

The expressions

$$\int_{-\infty}^a \phi(x)dx, \qquad \int_{-\infty}^\infty \phi(x)dx,$$

may be interpreted by analogy with the above, and may represent definite quantities or not, i.e., may exist or not, according to the character of the integrand. Obviously, if the x-axis is not an asymptote, these integrals do not exist.

Example 3:

$$\int_1^\infty \frac{dx}{x} = \lim_{x\to\infty} \Big[\log x\Big]_1^x = \lim_{x\to\infty} \log x.$$

Since this limit does not exist, the integral does not exist.

Example 4:

$$\int_{-\infty}^{0} e^x dx = \lim_{x\to-\infty} \Big[e^x\Big]_x^0 = 1.$$

5. **Expressibility of integrals.** We have seen in the paragraphs above that for any integrand continuous except at a finite number of points there exists an integral function continuous at every point excepting possibly points at which the integrand has a vertical asymptote. The fact does not depend upon the manner in which the integrand is expressed. The integral exists just as truly even if the integrand cannot be expressed in terms of the elementary algebraic and transcendental functions. But the integral functions can be found by the formulas of Chapter IV and the devices of the present chapter only when both integrand and integral may be expressed in terms of these elementary functions. Many important integrals can not be so expressed, and some types of them are treated in later chapters of the present work.

It should be observed that the integral may belong to a different class of functions from the integrand. For example

$$\int \frac{dx}{x-1} = \log(x-1) + c,$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \sin^{-1} x + c,$$

$$\int \tan x\, dx = \log \sec x + c,$$

$$\int \frac{dx}{1-x^2} = \tanh^{-1} x + c = \log \frac{1+x}{1-x} + c.$$

Thus the integrand may be algebraic and rational while the integral is transcendental, or the former may be algebraic and irrational and the latter transcendental, or the former transcendental and the latter transcendental of another type.

Now frequently very important integrals, where the integrand is expressed in terms of irrational algebraic functions or elementary

transcendental functions, serve to define functions of a class not expressible in terms of these elementary functions just as trigonometric and logarithmic functions are not expressible in terms of algebraic functions or in terms of each other. Such integrals may be plotted by certain suitable devices, some of which are to be studied later, and some of these integrals are sufficiently important to be given definite names and have their values published in tabular form, as are trigonometric, logarithmic, exponential and hyperbolic functions. Important examples are the **"probability integral"**

$$\int e^{-x^2} dx,$$

which is treated in all works on Probability, or Method of Least Squares and in most tables of integrals and physical tables; and **elliptic integrals** involving integrals of algebraic expressions with an indicated square root of a polynomial of the third or fourth degree. These latter integrals are the subject of an extensive theory, and their values are tabulated in most collections of physical tables and in most tables of integrals.

Another important function, called the **gamma function,** the Greek letter Γ being used as the function symbol, is defined by a definite integral, the independent variable of the function entering as a parameter in the integrand:

$$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1} dx \tag{4}$$

If we apply integration by parts, putting $u = e^{-x}$, $dv = x^{n-1}dx$ we obtain

$$\int_0^\infty e^{-x} x^{n-1} dx = \left[-\frac{1}{n} e^{-x} x^n \right]_0^\infty + \frac{1}{n} \int_0^\infty e^{-x} x^n dx,$$

and since the first term on the right is zero at both limits (see Art 46), we have

$$n\Gamma(n) = \Gamma(n+1) \tag{5}$$

From (4) we see that $\Gamma(1) = 1$, then from (5) that $\Gamma(2) = 1$, and for positive integral values of n we have

$$\Gamma(n) = (n-1)! \tag{6}$$

In mathematical and physical tables this function is tabulated for fractional values of n between 1 and 2, so that with these tables and (5) above we are able to find the value of the function for all positive values of n.

EXERCISES XXV

1. Show that the following integrals are continuous functions for all values of x. Sketch the graphs for both integrand and integral in each case:

(a) $\int x^{-\frac{1}{3}}\,dx$; (b) $\int \frac{dx}{\sqrt{4-x}}$.

2. Show that the integral in each of the following has one or more vertical asymptotes:

(a) $\int \frac{dx}{1-x^2}$; (b) $\int \tan x\,dx$; (c) $\int \frac{dx}{x^2}$.

3. Test the validity of the definite integral processes indicated below:

(a) $\int_0^5 \log x\,dx$; (b) $\int_0^{\pi} \tan x\,dx$; (c) $\int_0^{\infty} e^{-x}dx$; (d) $\int_0^2 \frac{dx}{(1-3x)^2}$

(e) $\int_1^{\infty} \frac{dx}{x^3}$; (f) $\int_{-\infty}^{-1} x^{-\frac{1}{3}}dx$.

4. Show that the area between a hyperbola and one of its asymptotes, on one side of its axis, is not finite.

Art. 29. Use of Tables of Integrals

While the devices of Articles 24–27, properly interpreted and skilfully applied, serve to integrate any function whose integral may be expressed in terms of elementary functions, yet in the interests of economy of time and labor the student is advised to become familiar with a table of integrals. Tables will differ in completeness and in arrangement, but a few remarks here should aid in effective use of any such works for reference.

1. **Classification of integrals.** Integrals are classified according to the nature of the integrand. First we have rational algebraic functions, arranged according to the nature of the factors of the denominator. Second we have irrational algebraic functions, grouped according to the nature of the radical present. Another group will include functions involving the elementary transcendental functions; and a final collection will usually give values of certain definite integrals, where the indefinite integral cannot be expressed in terms of elementary functions.

The four divisions indicated here may be variously subdivided, and each division will include a limited number of integral forms, selected with a view to economy of space, as well as to facility of expressing in terms of them the integrals, belonging to the given division, that arise

most frequently in practice. A reference table of integrals will be used most efficiently when familiarity with the plan and forms of the table is combined with a working knowledge of the devices of the present chapter.

2. **Preparation of integrand.** When an integral arises for which the value is not readily written, a form in the reference table may serve as a formula, without material transformation. But more often a transformation of some kind will be required to reduce to listed forms. In this the devices of the present chapter are found useful, and very frequently they will yield the value of the integral more readily without reference to the table than by use of its forms.

For treatment of rational algebraic functions, the method of Article 27 is entirely adequate to arrive at forms included in any reference table, and in many cases possible reductions by partial fractions may be omitted by use of such tables.

In treating irrational algebraic functions it is necessary to transform algebraically until the integrand is expressed as a sum of terms, each involving only one radical in the variable. This will require rationalization of the denominator, in case the denominator contains two or more terms, at least one of which is a radical. If a term consists of a rational fraction multiplied or divided by a radical, the rational fraction may be treated by the partial fraction process.

As to transcendental functions, very little can be said of general application. Familiarity with the table itself, and with the usual trigonometric and logarithmic transformations, is most advantageous. It may be observed, however, that a function of a function F[f(x)], where the function F is an elementary transcendental function, seldom occurs in reference tables, and when arising in practice should be treated first by the substitution $f(x) = u$. Success in integrating such expressions will depend on the character of other factors present. For example,

$$\int x^5 e^{x^2} dx = \frac{1}{2}\int u^2 e^u du = \frac{1}{2}e^u(u^2 - 2u + 2) + c,$$

where $u = x^2$. If the factor x^5, or some other odd power of x, had not been present, the indefinite integral could not have been expressed.

In the following miscellaneous exercises in integration the student should give attention to selection of method, making use of any reference tables available.

EXERCISES XXVI

1. $\int (e^{5x} + a^{5x})dx$;

2. $\int \frac{(a^x - b^x)^2 dx}{a^x b^x}$;

3. $\int \left(\cos\frac{\theta}{3} - \sin 3\theta\right)d\theta$;

4. $\int \cos(\log x)\frac{dx}{x}$;

5. $\int (\sec\phi - \tan\phi)^2 d\phi$;

6. $\int \frac{ax\,dx}{x^4 + e^4}$;

7. $\int \frac{du}{\sqrt{a^2 - (u+b)^2}}$;

8. $\int \frac{ax\,dx}{x^4 - e^4}$;

9. $\int \frac{dx}{x\sqrt{c^2x^2 - a^2b^2}}$;

10. $\int \tan^4\theta \sec^4\theta\, d\theta$;

11. $\int \frac{(x^2 + x - 1)}{x^3 + x^2 - 6x}dx$;

12. $\int \frac{(2z^2 - 5)dz}{z^4 - 5z^2 + 6}$;

13. $\int \frac{(3x+2)dx}{x(x+1)^3}$;

14. $\int \frac{aS^2dS}{(S+a)^3}$;

15. $\int \frac{4dx}{x^4 + 1}$;

16. $\int \frac{x^3 + x - 1}{(x^2+2)^2}dx$;

17. $\int \frac{x^{\frac{1}{2}}\,dx}{x^{\frac{1}{2}} + 1}$;

18. $\int \frac{dx}{x(x+1)^{\frac{1}{2}}}$;

19. $\int \alpha \sin\alpha\, d\alpha$;

20. $\int x\tan^{-1}x\,dx$;

21. $\int \cos\theta \log\sin\theta\, d\theta$;

22. $\int \frac{dS}{(a^2 + S^2)^3}$;

23. $\int \frac{r^8 dr}{\sqrt{1 - r^3}}$;

24. $\int_{1.5}^{2} \frac{dx}{2x - 3}$;

25. $\int_{1.5}^{2} \frac{dx}{\sqrt{2x - 3}}$;

26 $\int_{o}^{a} \frac{x^2dx}{\sqrt{ax - x^2}}$;

27. $\int_{-1}^{+1} \frac{dx}{x^2 + 5x + 4}$;

28. $\int_{o}^{\infty} e^{-2x}dx$;

29. $\int \frac{x^5dx}{\sqrt{2ax - x^2}}$;

30. $\int \frac{(x + \sin x)dx}{1 + \cos x}$;

31. $\int_0^{\frac{\pi}{6}} \frac{d\theta}{\sqrt{1 - 1/4 \sin^2\theta}}$;

32. $\int e^{2x} \tan^3 x \, dx$;

33. $\int_0^{\infty} \frac{\cos 3\alpha}{1 + 9\alpha^2} \, d\alpha$;

34. $\int_0^1 e^{-x^2} \, dx$;

35. $\int \frac{dx}{\sqrt{a + 2e^{2x}}}$;

36. $\int_0^{3a} \frac{2x \, dx}{(x^2 - a^2)^{\frac{2}{3}}}$;

37. $\int_0^1 \frac{dx}{e^x + e^{-x}}$;

38. $\int_0^{\log 2} \sqrt{e^x - 1}$; dx;

39. $\int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{2 - \sin^2\theta}}$;

40. $\int y \cosh 3y \, dy$;

41. $\int \sinh 3x \cosh 3x \, dx$;

42. $\int e^{\frac{x}{2}} \sin x \, dx$;

43. $\int_0^1 \frac{\log x}{1 - x} dx$;

44. $\int \frac{(1 + 3x)dx}{(2 + 3x + 5x^2)^{\frac{3}{2}}}$;

45. $\int \frac{dx}{3x\sqrt{5 - 6x}}$;

46. $\int_0^{\infty} x^{\frac{5}{2}} e^{-x} \, dx$;

47. $\int_0^{\frac{1}{2}} \frac{dx}{\sqrt{(1 - x^2)(4 - x^2)}}$;

48. $\int_0^{\infty} e^{-4x^2} \, dx$.

BOOKS FOR SUPPLEMENTARY READING

A Short Table of Integrals, B. O. Peirce (Ginn, 1910).
Advanced Calculus, W. F. Osgood (Macmillan, 1925), Chapters I, II, IX and XIX.
Advanced Calculus, E. B. Wilson (Ginn, 1912), Chapters XI, XIII and XIV.
Integral Calculus, W. E. Byerly (Ginn, 1888).
Mathematical Analysis, Goursat-Hedrick Vol. I (Ginn, 1904), Chapters IV and V.

CHAPTER VII

CERTAIN STANDARD CURVES, EQUATIONS, TRANSFORMATIONS

Art. 30. Certain Uses of Coordinates

In the preceding chapters we have been dealing with functions, and we have found the graphs of functions very useful. The graphs we have considered have been a means of expressing the values of a function $y = f(x)$ as ordinates, measured vertically in some convenient scale, corresponding to values of x on a range laid off horizontally to some scale. The ordinate, y, and the abscissa, x, are the coordinates of a point on the graph, and the graph is the locus of points whose coordinates have the relation $y = f(x)$. Our interest in graphs has not been geometrical, but has pertained to their use in interpretation of the functional relation.

In the present chapter attention is given to geometrical properties of certain plane figures, regarded as loci of points whose coordinates are related by certain equations. We shall confine our attention chiefly to figures and equations which are so thoroughly standardized and established in mathematical literature that a familiarity with them is an essential part of the equipment of any student of mathematical subjects.

The position of a point in a plane or on a given surface may be determined by two quantities, called the **coordinates** of the point. Various systems of coordinates are in use. The latitude and longitude of a place on the surface of the earth determine its position; and latitude and longitude, together with established "zero" lines, the equator and the prime meridian, constitute a system of coordinates on the earth's surface.

If in a plane a certain zero point or origin be fixed as O (Fig. 33), and two lines be drawn through O, fixed in both direction and sense as Ox and Oy, then O may be joined to any point P in the plane by two straight line segments, one in each of the two fixed directions, as OD and DP. The lengths of these line segments, positive or negative according to their agreement or disagreement in sense with the fixed lines, may be taken as coordinates of the point P. With a fixed unit of length, such a system constitutes a **Cartesian** system of coordinates,

so called after the French philosopher Descartes, who first used them successfully.

The most commonly used system of coordinates is the **rectangular** Cartesian system, the line Ox, the axis of x, being horizontal with the positive sense to the right, and the line Oy, the axis of y, being vertical with the positive sense upward. The interest being mainly geometrical, the coordinates x and y are measured with the same linear unit. To each pair of values, (x, y), there corresponds a definite point P in the plane, and for any point P in the plane there are definite coordinates x and y.

Another system of coordinates much in use is the **polar** system. A fixed point O is taken as the **pole,** and a straight line CA through O, fixed in direction and sense, usually with the positive sense to the right, is taken as **initial line.** Given any point P in the plane (Fig. 34), a line joining O to P is called the **radius vector** of the point P, and its

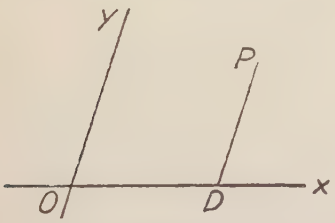


Fig. 33

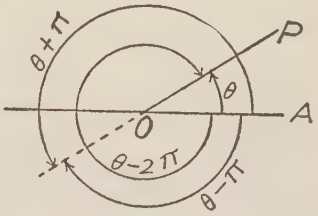


Fig. 34

length is denoted by ρ, positive if the positive sense of the radius vector is regarded as from O to P, otherwise negative. The angle between OA and the positive end of the radius vector, positive when measured counter clockwise from OA, is called the **vectorial angle,** and is denoted by θ. The quantities ρ and θ constitute the coordinates of the point P. Corresponding to any pair of coordinates (ρ, θ), there is a definite point P, but a given point P may be denoted with the equal exactness by an infinitude of different coordinate values. If (ρ, θ) may determine a point P, then this same point is given by $(\rho, \theta \pm 2n\pi)$ and by $[(-\rho, \theta \pm (2n + 1)\pi]$ where n is any positive integer.

1. **Distance between points.** Let P_1 and P_2 be two points in a plane, and let their coordinates in a rectangular system be respectively (x_1, y_1) and (x_2, y_2). The line joining P_1 to P_2 is the diagonal of a rectangle (Fig. 35) whose edges are $x_1 - x_2$ and $y_1 - y_2$. The distance, d, between the points is therefore

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \qquad (1)$$

Again if (ρ_1, θ_1) and (ρ_2, θ_2) are the polar coordinates of these points, the side P_1P_2 of the triangle OP_1P_2, has its length given by

$$d = \sqrt{\rho_1^2 + \rho_2^2 - 2\rho_1\rho_2 \cos(\theta_1 - \theta_2)}. \tag{2}$$

2. **Areas of polygons.** Consider a triangle whose vertices are P_1, P_2 and P_3, any three points in the plane. The line joining two of these points, the ordinates of the two points and the intercepted portion of the x-axis bound a trapezoid as $P_1D_1D_2P_2$ (Fig. 36). The area of the triangle may be expressed in terms of three such trapezoids, as $A = -P_1D_1D_2P_2 - P_2D_2D_3P_3 + P_1D_1D_3P_3$, and expressing hese areas in terms of rectangular coordinates of the points we have

$$A = \frac{1}{2}(x_1 - x_2)(y_1 + y_2) + \frac{1}{2}(x_2 - x_3)(y_2 + y_3) + \frac{1}{2}(x_3 - x_1)(y_3 + y_1)$$

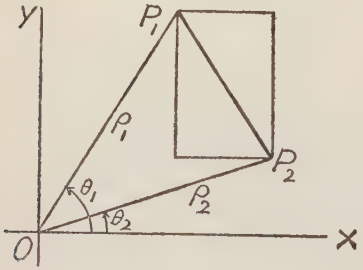


Fig. 35

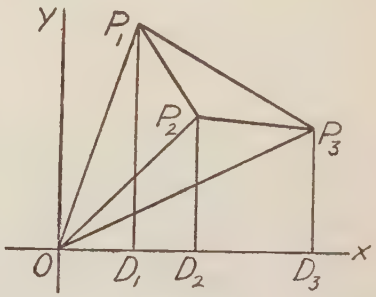


Fig. 36

and this reduces to

$$A = \frac{1}{2}(x_1y_2 + x_2y_3 + x_3y_1 - x_2y_1 - x_3y_2 - x_1y_3). \tag{3}$$

This result is conveniently written in the form of a determinant,

$$A = \frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}. \tag{3'}$$

It will be noticed that the three vertices are numbered in counter clockwise order around the triangle. To number these points in reverse order would have the effect of interchanging two rows of the

determinant, and would therefore, change the sign, but not the numerical value of the quantity.

In similar manner the area of the triangle may be expressed in terms of three triangles, as $A = -P_1OP_2 - P_2OP_3 + P_1OP_3$, and these areas may be expressed in terms of the polar coordinates of the vertices

$$A = \frac{1}{2}\rho_1\rho_2\sin(\theta_2 - \theta_1) + \frac{1}{2}\rho_2\rho_3\sin(\theta_3 - \theta_2) + \frac{1}{2}\rho_3\rho_1\sin(\theta_1 - \theta_3). \quad (4)$$

Here, again, a reversal of the order of numbering would change the sign of the quantity.

Now let the vertices of a polygon, taken in counter clockwise order around the figure, be P_1, P_2 . . . . P_n. The polygon may be divided into triangles by joining P_1 to each of the other vertices. Writing the expression (3) for each of these triangles, then adding, we obtain for the **area of the polygon**

$$\begin{aligned} A &= \frac{1}{2}[(x_1 - x_2)(y_1 + y_2) + (x_2 - x_3)(y_2 + y_3) + \ldots \\ &\quad + (x_n - x_1)(y_n + y_1)] \qquad (5) \\ &= \frac{1}{2}(x_1y_2 + x_2y_3 + \ldots + x_ny_1 - x_2y_1 - \ldots - x_1y_n). \end{aligned}$$

This may be remembered by association with the array

$$\begin{array}{ccccccccc} x_1 & x_2 & x_3 & \ldots\ldots\ldots\ldots & x_n & x_1 \\ \times & \times & \times & & \times & \\ y_1 & y_2 & y_3 & \ldots\ldots\ldots\ldots & y_n & y_1 \end{array}$$

If each of the triangles into which the polygon is divided is expressed by (4) the sum becomes

$$\begin{aligned} A = \frac{1}{2}[&\rho_1\rho_2\sin(\theta_2 - \theta_1) + \rho_2\rho_3\sin(\theta_3 - \theta_2) + \ldots \\ &+ \rho_n\rho_1\sin(\theta_1 - \theta_n)]. \qquad (6) \end{aligned}$$

3. **Equation and locus.** A curve may be regarded as the locus of a point, i.e., it contains all points of the plane which fulfill a certain condition, and it contains no points which do not fulfill this condition. This condition may be capable of expression by means of an equation giving a relation between the coordinates of the point. Such an equation is said to be the equation of the curve. One fundamental problem

of analytic geometry is to write the equation of a curve when the definition is known.

For example, let it be required to find the equation of a circle of radius 5, with center at (2, −7). This circle is the locus of points at distance 5 from (2, −7), therefore the coordinates (x, y) of any point on the circle must have the relation

$$\sqrt{(x-2)^2+(y+7)^2}=5.$$

Or

$$x^2+y^2-4x+14y+28=0.$$

This is the rectangular equation of the circle. The polar equation may be written

$$\rho^2+\rho_1^2-2\rho\rho_1\cos(\theta-\theta_1)=25,$$

where (ρ_1, θ_1) is the polar representation of the point (2, −7), and if the pole is taken at the origin and the initial line as the x-axis,

$$\rho_1^2=53, \quad \tan\theta_1=-\frac{7}{2}.$$

Again, let it be required to find the rectangular equation of a straight line through the point (3, 2) parallel to the line joining the points (−5, 12) and (4, 3). Any point (x, y) on the required line, with the last two given points, must form a triangle whose area is equal to the area of the triangle of the three given points. Using (3′) then, we have

$$\begin{vmatrix} x & y & 1 \\ -5 & 12 & 1 \\ 4 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 3 & 2 & 1 \\ -5 & 12 & 1 \\ 4 & 3 & 1 \end{vmatrix},$$

or, after expanding and reducing, $x+y=5$.

Another fundamental problem of analytic geometry is to determine, or draw, the curve whose equation is given. Such a curve is called the **locus of the equation.** Here the methods of Chapter I are applicable, but we consider additional devices which pertain to the geometrical character of the locus, and which may be used to reduce the labor of plotting the locus.

(a) **Symmetry.** A curve is **symmetrical with respect to a line** if every chord of the curve drawn perpendicular to the line is bisected by

the line. A curve is **symmetrical with respect to a point if every** chord of the curve drawn through the point is bisected at the point.

The following facts should be evident:

A curve is symmetrical with respect to the x-axis if only even powers of y occur in its equation; i.e., if y may be replaced by $-y$ without changing the equation.

A curve is symmetrical with respect to the y-axis if only even powers of x occur in its equation.

A curve is symmetrical with respect to the origin if all terms of its equation are of even degree in x and y, or all of odd degree in x and y. A constant term may be regarded as of even degree.

A curve is symmetrical with respect to the pole if only even powers of ρ occur in its polar equation, or if ρ is a periodic function of θ with period $\frac{\pi}{n}$, where n is any integer.

(b) **Bounds of a locus.** If the equation of a curve be solved for one variable in terms of the other, say for y in terms of x, it may appear that real values of y occur only for certain restricted values of x. In this way, solving first for y and then for x, we may be able to fix bounds for the locus. The same treatment may be applied to the polar equation of a curve.

(c) **Asymptotes.** It is not desirable to discuss at this time the determination of asymptotes in general, but it should be helpful to consider the following cases:

If y becomes indefinitely large, positive or negative, as x approaches a certain constant, a, then there is a vertical asymptote to the curve at $x = a$. This will evidently be the case if y is expressed as a function of x by a fraction whose denominator contains the factor $x - a$.

If x becomes indefinitely large as y approaches b, then there is a horizontal asymptote at $y = b$.

For example, let it be required to draw the curve whose equation is

$$x^2y^2 - y^2 - x^2 = 0.$$

The curve is symmetrical with respect to both axes and with respect to the origin, so that when we have completed that portion of the curve in the first quadrant the other parts may be drawn by symmetry. We have

$$y^2 = \frac{x^2}{x^2 - 1},$$

which shows that y is imaginary if x is numerically less than unity, so that the curve lies to the right of the dotted line at $x = 1$ (Fig. 37). Also, since

$$x^2 = \frac{y^2}{y^2 - 1},$$

the curve lies above the line at $y = 1$. As x approaches unity y increases indefinitely, so that a vertical asymptote is at $x = 1$. Simi-

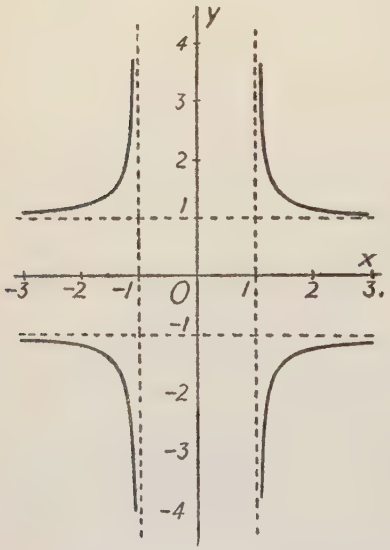


Fig. 37

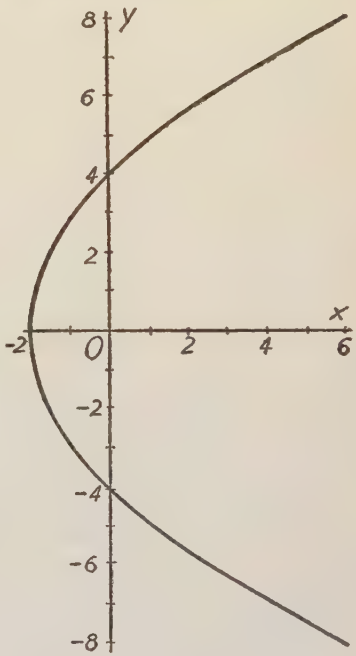


Fig. 38

larly there is a horizontal asymptote at $y = 1$. Three points on the curve are $\left(2, \frac{2}{\sqrt{3}}\right)$ $(\sqrt{2}, \sqrt{2})$ and $\left(\frac{2}{\sqrt{3}}, 2\right)$, and with this information the curve may be drawn.

Again, let us draw the locus of $y^2 = 8x + 16$. Here we have symmetry with respect to the x-axis only. We shall have y imaginary if $x < -2$. Points on the curve are $(-2, 0)$, $(0, 4)$ $(6, 8)$, and from this the curve may be drawn (Fig. 38).

4. **Intersections and parameters.** Given two equations, each repre-

senting a locus, any values of the variables that satisfy both equations are the coordinates of a point common to the two curves. Thus, to find the points of intersection of two curves, we have only to regard their equations as simultaneous and solve for x and y, or for ρ and θ, according to the system of coordinates used. In the process of solving the simultaneous equations various combinations are made to form new equations. It should be noted that these new equations also represent curves going through the points of intersection of the original curves. Usually the solution consists in finally arriving at systems of lines through the required points parallel to the coordinate axes. For instance, to find the intersections of the circle $x^2 + y^2 = 25$ with the line $x - y = 1$, we may substitute $x = y + 1$ in the former, obtain $2y^2 + 2y = 24$ or $y^2 + y - 12 = 0$, which is equivalent to the two horizontal lines $y = 3$, $y = -4$, and using these values, with the second equation we have the two vertical lines $x = 4$, $x = -3$. These two systems of lines determine the required points, but also two others. The required points are (4, 3) (−3, −4). The other two points, (−3, 3) and (4, −4) are not on the given circle and line.

If $f(x, y) = 0$, and $\phi(x, y) = 0$ are equations of two loci, then

$$\phi(x, y) + kf(x, y) = 0. \tag{7}$$

where k is any constant, is the equation of a new curve passing through all points common to the given loci. Different values of k will give different members of a set or **family** of curves, all of which intersect in the common points mentioned. The quantity k used here is an **arbitrary constant,** when thinking of a single curve of the family, but when thinking of the family as a whole k is regarded as assuming a whole series of values and is called a **parameter.** The word parameter is used in analytic geometry in two distinct relations, as shown in the following:

(a) A single equation in x and y or ρ and θ may involve one or more arbitrary constants, denoted by letters. For each set of values of these constants a definite curve is determined, but different curves are obtained by assigning different values to the constants. The family of curves so determined is called a **one-parameter family,** a **two-parameter family, three-parameter family,** etc., according to the number of quantities considered as parameters. Here the parameter does not vary with x and y or ρ and θ, but is independent, and a definite relation is given to the two coordinates only when every parameter is assigned a fixed value.

(b) In previous chapters we have made use of auxiliary variables, a special case of which occurs in the use of parametric equations, as

$x = f(t)$, $y = \phi(t)$. Here the parameter, t, is a third variable, which must vary with x and y, and different values of t do not pertain to different curves, but to different points of the same curve. Elimination of t gives the equation of the curve, free from the parameter. Such parametric equations may also involve arbitrary constants, as $x = a \cos t$, $y = a \sin t$. Here a is a parameter in the sense of arbitrary constant, while t is a parameter in the sense of auxiliary variable. Squaring and adding these equations gives $x^2 + y^2 = a^2$, a family of concentric circles, a being the parameter of the family. If, however, we choose to regard a as auxiliary variable and t as arbitrary constant then we obtain $y = x \tan t$, a family of lines through the origin, t, the angle made with the x-axis, being the parameter of the family.

EXERCISES XXVII

1. Show that the triangle whose vertices are (−2, 4) (5, 3) and (2, 7) is isosceles, find its base and altitude, and thence its area. Also find the area by formula (3′).

2. Show by lengths of sides and diagonals, that (1, 2), (13, −3), (16, 1) and (4, 6) are the vertices of a parallelogram. Find the area by formula (5).

3. The polar coordinates of two points are $\left(3, \frac{\pi}{4}\right)$ and $\left(8, \frac{5\pi}{12}\right)$. Find the distance between these points. By an application of formula (4) write the polar equation of a line which is parallel to the line joining these points, but at a distance 3 from it.

4. Write the equation of the perpendicular bisector of the line joining (−3, 2) to (9, 7), considered as locus of points equidistant from these points.

5. For what relation between x and y will the area of the triangle (2, 7) (4, 3) (x, y) be zero? What is the locus of the point (x, y) so conditioned?

6. Write the equation of the locus of a point whose distance from (0, 4) is equal to four times its ordinate. Draw the curve.

7. Work problem 6 in polar coordinates, taking the given point as $\left(4, \frac{\pi}{2}\right)$ and observing that the ordinate is $\rho \sin \theta$.

8. Write the equation of the locus of a point the sum of whose distances from (3, 0) and (−3, 0) is 10. Discuss the locus with regard to symmetry and bounds. Draw the curve.

9. A point, (x, y), is so placed that the line joining it to (2, 3) is the diagonal of a rectangle whose area is 8. If one side of the rectangle

is always horizontal, what is the locus of the point (x, y)? Write the equation and draw the curve.

10. Write the equation of the locus of a point at distance 6 from the point $(-4, 4)$. Draw the curve.

11. Work problem 10 in polar coordinates, observing that the given point is $\left(4\sqrt{2}, 3\frac{\pi}{4}\right)$.

12. Find the equation of the locus of points twice as far from (6, 0) as from the origin. Show that this is a circle of radius 4, center at $(-2, 0)$.

13. Work problem 12 in polar coordinates.

14. A point moves so that its distance from the point (4, 0) is k times its distance from the point $(-4, 0)$. Write the equation of the locus, using k as an arbitrary constant. Draw the locus for several values of k, as $\frac{1}{2}$, 1, 2, 3, etc., and show that each is a circle.

15. Write the equation of a one-parameter family of circles passing through the points common to $x^2 + y^2 + 6y = 16$ and $x^2 + y^2 - 6y = 16$. Draw several curves of this family. Show that the common points are (4, 0) and $(-4, 0)$.

16. Show by careful drawing that curves in 14 intersect curves in 15 at right angles.

17. Draw the curve determined by the parametric equations $x = 4 \sec t$, $y = 3 \tan t$. Show that this curve is the locus of points the difference of whose distances from (5, 0) and $(-5, 0)$ is 8.

18. Discuss the two-parameter family of curves whose parametric equations are $x = a \cos t$, $y = b \sin t$, a and b being the family parameters and t the auxiliary variable.

Art. 31. The Straight Line

Given any straight line in the plane, and a definite system of coordinates, there are certain constants, or parameters, definitely associated with the line. The names and notations of these parameters are fairly well standardized in mathematical literature. They are: The x-**intercept,** a; the y-**intercept,** b; the tangent of the angle made with the x-axis, called the **slope** of the line, m; the perpendicular distance from the origin, p; and the angle that this perpendicular makes with the x-axis, α (Fig. 39).

It is evident that $m = \tan\left(\alpha \pm \frac{\pi}{2}\right)$, so that either α or m may be

found from the other. Excepting this relation, any two of these parameters are independent of each other, and any two independent parameters may be used to determine the line. The following equations, each using a pair of these parameters are standard forms for the equation of a straight line.

The **slope-intercept,** or simply the **slope** equation,

$$y = mx + b, \tag{1}$$

is evidently true for any point (x, y) on the line.

The **intercept** equation is

$$\frac{x}{a} + \frac{y}{b} = 1, \tag{2}$$

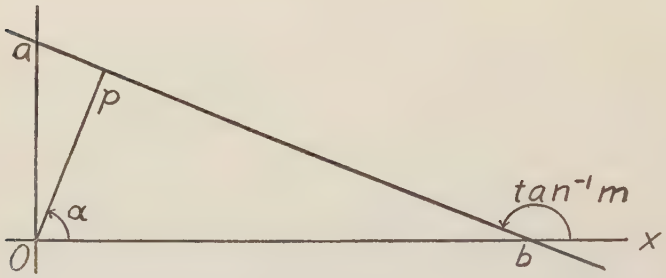


Fig. 39

and it is evident that a and b are the intercepts of this locus on the x- and y-axis; furthermore, this equation is a result of the slope-intercept equation, since $m = -\frac{b}{a}$.

The **normal** equation is

$$x \cos \alpha + y \sin \alpha = p, \tag{3}$$

and may be derived from the intercept equation by observing that $a = p \sec \alpha$ and $b = p \csc \alpha$.

The standard **polar** equation of a line is

$$\rho \cos (\theta - \alpha) = p. \tag{4}$$

That any point (ρ, θ) on the line must satisfy this equation is evident geometrically, and the equation is at once deduced from the normal equation by putting $x = \rho \cos \theta$ and $y = \rho \sin \theta$.

In general, an equation of the first degree in x and y may be reduced to the slope, intercept and normal forms, and therefore represents a straight line. For example, the equation $5x - 12y + 6 = 0$ may be written in the slope form,

$$y = \frac{5x}{12} + \frac{1}{2},$$

the equation being solved for y and showing that $m = \frac{5}{12}$ and $b = \frac{1}{2}$. Again we may write

$$\frac{x}{-\frac{6}{5}} + \frac{y}{\frac{1}{2}} = 1,$$

characterized by unity on the right, and showing $a = -\frac{6}{5}$, $b = \frac{1}{2}$. Finally, observing that $5^2 + 12^2 = 13^2$, we may divide by -13, obtaining

$$-\frac{5}{13}x + \frac{12}{13}y = \frac{6}{13}.$$

This equation is characterized by the fact that the sum of the squares of the coefficients of x and y is unity, and therefore they become the cosine and sine of some angle. Here $\cos \alpha = -\frac{5}{13}$, and $\sin \alpha = \frac{12}{13}$, $p = \frac{6}{13}$. Usually, in numerical cases, the equation is written so that p is positive, but this is not essential. The parameters p and α are in reality polar coordinates of the foot of the perpendicular on the line, and are properly subject to the ambiguities that pertain to the polar coordinates of a point. Changing signs in the last equation would give $\cos \alpha = \frac{5}{13}$, $\sin \alpha = -\frac{12}{13}$, $p = -\frac{6}{13}$, and these values are as effective as the former values.

Three special cases must be noted. First, a line parallel to the y-axis has an equation $x = c$, and, since it does not contain y, it cannot be written in the slope form. Second, a line through the origin has no constant term in its rectangular equation, and therefore cannot be written in the intercept form. Third, in polar coordinates, the equation of a line through the pole takes the form $\theta = \alpha \pm \frac{\pi}{2}$.

Except for the ambiguities in p and α, the forms discussed above are each unique for a given line. Each has certain advantages, and the choice of form depends on the character of information by which a line is determined, and on the nature of information desired. Two independent conditions determine a line, hence determine all its parameters. But often it is convenient to write the equation of a line without using its standard parameters. Two standard forms are in use for this purpose.

The **point-slope** equation is

$$y - y_1 = m(x - x_1). \tag{5}$$

Here (x_1, y_1) is obviously a point on the line, and the equation may be written if the slope and any point on the line are known.

If two points on the line are known we may use the **two-point** equation.

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1} \tag{6}$$

This is a linear equation in x and y, and it is obviously satisfied if $x = x_1$ and $y = y_1$, or if $x = x_2$ and $y = y_2$, therefore it represents a line through the points (x_1, y_1) and (x_2, y_2). Another form of the two point equation is

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0, \tag{6'}$$

which expresses the condition that the area of the triangle determined by (x, y) and the two given points is zero.

To find the **angle that a line** AB **makes with a line** CD we observe that this angle is the difference between the angles that the lines make with the x-axis, $\phi = \theta_1 - \theta_2$ (Fig. 40). Now if the slopes of AB and CD are m_1 and m_2 respectively, then

$$\tan \phi = \frac{\tan \theta_1 - \tan \theta_2}{1 + \tan \theta_1 \tan \theta_2} = \frac{m_1 - m_2}{1 + m_1 m_2}. \tag{7}$$

If on the other hand, one line, as CD, is given and we wish to determine **the slope of** AB **so that it shall make a given angle** ϕ **with** CD, we have only to solve this equation for m_1, obtaining

$$m_1 = \frac{m_2 + \tan \phi}{1 - m_2 \tan \phi}. \tag{8}$$

If AB and CD are **perpendicular,** then tan ϕ is infinite, and $1 + m_1m_2 = 0$, or $m_1 = -1/m_2$.

The normal equation has an important use as a distance formula. In the equation $x \cos \alpha + y \sin \alpha - p = 0$, the quantity p represents the distance from the origin to the line, and the equation $x \cos \alpha + y \sin \alpha - p = d$ represents a line parallel to the first whose distance from the origin is $p + d$, therefore, any point on the second line is at distance d from the first line. The **distance of any point** (x_1, y_1) **from a line** $x \cos \alpha + y \sin \alpha = p$ is, therefore

$$d = x_1 \cos \alpha + y_1 \sin \alpha - p. \tag{9}$$

This distance formula furnishes us a convenient means of interpreting certain equations. We illustrate by two examples.

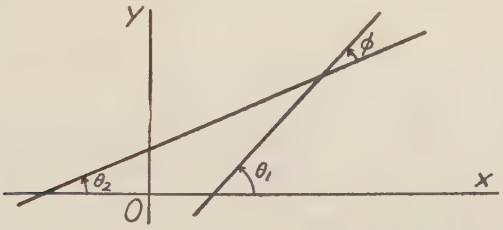


Fig. 40

Consider the two lines, $3x + 5y = 9$, and $x - 7y = 4$. The equation

$$(3x + 5y - 9) + k(x - 7y - 4) = 0,$$

for any value of k, represents a line through the intersection of the given lines. In particular, if k is the ratio of the normalizing factors of the two lines, i.e.,

$$\pm k = \frac{\sqrt{3^2 + 5^2}}{\sqrt{1^2 + 7^2}} = \frac{\sqrt{34}}{5\sqrt{2}},$$

then we may write

$$\frac{3x + 5y - 9}{\sqrt{34}} = \pm \frac{x - 7y - 4}{5\sqrt{2}},$$

and, each side being a distance formula, the equation represents the locus of points equidistant from the two lines, that is, the two angle bisectors.

Again, the equation

$$(3x + 5y - 9)(x - 7y - 4) = k$$

may be written

$$\frac{(3x + 5y - 9)}{\sqrt{34}} \frac{(x - 7y - 4)}{5\sqrt{2}} = \frac{k}{10\sqrt{17}},$$

and therefore represents the locus of points the product of whose distances from the two given lines is the constant on the right in the last equation. The product of distances being constant, as one distance becomes very small the other becomes very large, so that each line is an asymptote to the curve.

EXERCISES XXVIII

1. Write the rectangular equations of lines determined as follows:

(a) The slope is 3, the x-intercept is 5.

(b) The y-intercept is -2, passes through (7, 3).

(c) Passes through (2, 9) parallel to $3x - 8y = 7$.

(d) Passes through (1, 3) and $(-2, 5)$.

(e) Slope is $\frac{3}{4}$, distance from origin is 8.

(f) Through (1, 8), at distance 2 from $(-3, 4)$.

2. Write the polar equation of a line

(a) Perpendicular to the initial line, 2 units from the pole.

(b) Distance 5 from origin, passes through $\left(10, \frac{\pi}{4}\right)$.

(c) Through the pole, perpendicular to

$$\rho \cos\left(\theta - \frac{\pi}{3}\right) = 6.$$

3. Consider as the base of a triangle the line joining (1, 5) to (8, -2), and as the vertex the point (7, 6).

(a) Find the length of the base.

(b) Find the equation of the line of which the base is a segment.

(c) Find the altitude, by a distance formula.

(d) Find the area, as half of base by altitude.

(e) Find area by formula, Art. 30.

4. Find tangents of the angles of the triangle in 3.

5. Prove that the following three lines meet in a point: $3x + 4y = 7$, $2x - 3y = 5$, $x - 10y = 3$. (Hint: Write the determinant of coefficients and constant terms.)

6. The three sides of a triangle are $4x - 3y = -10$, $4x + 3y = 15$, $5x + 12y = 26$. Prove that the bisectors of the angles of the triangle meet in a point.

7. A line has the equation $3x + ky = 5$. Find in terms of k its slope, intercepts, and distance from the origin. Assign values to k that make the line fulfill the following conditions:

(a) Parallel to $7x + 4y + 5 = 0$.
(b) Makes an angle of 45° with the line $x - 2y = 6$.
(c) Distant one unit from the origin.
(d) Distant 2 from the point (2, 4).
(e) Through (7, 2).

8. The equation of a line is $x + ky = c$. Find k and c so that the line will

(a) Pass through (2, 4) and (−3, 7).
(b) Pass through (1, 5) perpendicular to $3x - 8y = 6$.
(c) Pass through the origin and at distance 2 from (5, −6).

9. Write the equation of a line through the point common to $x - 3y + 4 = 0$ and $5x + 7y = 1$, and also through (1, −3).

10. Giving attention to the triangle of problem 6, sketch the curves represented by

(a) $(4x - 3y + 10)(4x + 3y - 15) = 50$.
(b) $(4x - 3y + 10)(4x + 3y - 15)(5x + 12y - 26) = -65$.
(c) $(4x - 3y + 10)(4x + 3y - 15) + (5x + 12y - 26) = 0$.
(d) $(4x - 3y + 10)^2 = 5x + 12y - 26$.

11. Find three asymptotes to the curve $x^3 - xy^2 = 12$, and sketch the curve.

Art. 32. Transformation of Coordinates

We have seen that the equation of a locus or curve in rectangular coordinates has a form entirely different from the polar equation of the same locus. But even in rectangular coordinates the form of the equation of a given curve will depend on the choice of axes as well as on the character of the curve. It is often desirable, when given an equation in one system of coordinates, to determine the equation of the same locus in a new system of coordinates bearing a definite relation to the old. This is accomplished by expressing the old coordinates

of a general point in terms of the new coordinates of this point, and substituting these expressions in the given equation, thus obtaining an equation in the new coordinates.

We have already made use, in certain instances, of the relations between rectangular and polar coordinates. If the pole is taken at the origin and the initial line on the x-axis we have, for **transformation from rectangular to polar coordinates** (Fig. 41),

$$x = \rho \cos \theta, \quad y = \rho \sin \theta. \tag{1}$$

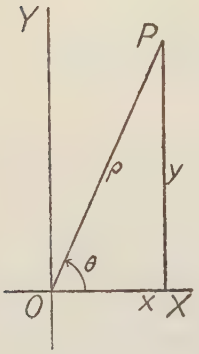


Fig. 41

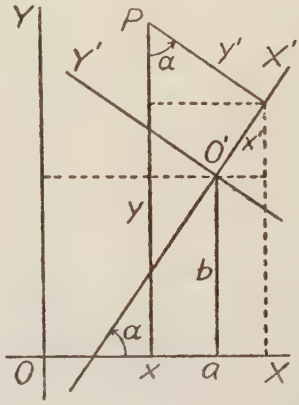


Fig. 42

To **transform from polar to rectangular coordinates** we have, by solving the above equations for ρ and θ, or from the figure directly,

$$\rho = \sqrt{x^2 + y^2}, \qquad \theta = \tan^{-1} \frac{y}{x}. \tag{2}$$

Auxiliary relations often useful in this process are:

$$\sin \theta = \frac{y}{\sqrt{x^2 + y^2}}, \qquad \cos \theta = \frac{x}{\sqrt{x^2 + y^2}}. \tag{2'}$$

A very important case of **transformation is from one rectangular system to another.** Let OX and OY (Fig. 42) be the old coordinate axes. If (a, b) is the point O′ with reference to the old axes, and if O′X′ makes angle α with OX, then, for any point P, whose new coordinates are (x′y′) and old coordinates are (x, y), we have

$$\begin{aligned} x &= a + x' \cos \alpha - y' \sin \alpha, \\ y &= b + x' \sin \alpha + y' \cos \alpha. \end{aligned} \tag{3}$$

by which we may transform any equation from the old system to the new. These relations are evident from the geometry of the figure, but they also result from an application of the distance formula, since

$$x' \cos \alpha - y' \sin \alpha + a = 0$$

is the equation of OY with reference to the new axes, and

$$x' \sin \alpha + y' \cos \alpha + b = 0$$

is the equation of OX with respect to the new axes.

If $\alpha = 0$, the transformation (3) is to new axes parallel to the old, and is known as **translation.** The equations become

$$x = x' + a, \qquad y = y' + b. \tag{4}$$

If $a = 0$ and $b = 0$, so that O′ coincides with O, then the transformation (3) is called **rotation** of axes, and the equations become

$$x = x' \cos \alpha - y' \sin \alpha,$$

$$y = x' \sin \alpha + y' \cos \alpha. \tag{5}$$

Any transformation by (3) may be accomplished by a translation and a rotation, and the two simple processes are usually preferable to the use of equations (3).

Since equations (3) substitute linear expressions for x and y, it is clear that a **new choice of rectangular axes cannot change the degree of an equation.**

From equations (4) we see that **translation does not change the terms of the equation that are of highest degree in x and y.**

Since equations (5) express x and y as homogeneous linear functions of x′ and y′ we see that **in rotation each term of the old equation gives only terms of the same degree in the new equation.**

These three facts indicate that, while a transformation of coordinates may change the equation of a curve, certain features of the equation persist under any transformations of translation and rotation. A convenient choice of axes, however, may have much to do with the simplification of an equation, and to determine new axes so that the new equation may possess certain characteristics is one of our important problems.

Given the equation $\rho = \sec^2 \frac{\theta}{2}$, let us transform to rectangular co-

ordinates, then choose new rectangular axes that simplify the equation as much as possible. We write

$$\rho = \frac{1}{\cos^2 \frac{\theta}{2}} = \frac{2}{1 + \cos \theta}$$

and using equations (2) and (2′) we have

$$\sqrt{x^2 + y^2} = \frac{2}{1 + \frac{x}{\sqrt{x^2 + y^2}}} = \frac{2\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2} + x}.$$

Rationalization of this yields the equation

$$y^2 = -4x + 4.$$

Now take $x = x' + 1$, $y = y'$, and we have

$$y'^2 = -4x',$$

and neither translation nor rotation can reduce the equation to simpler form. The student should plot the curve from the polar equation, draw both sets of rectangular axes, and check by use of rectangular coordinates.

Again, take the equation

$$3x^2 - 7xy + 2y^2 + 9x + 2y - 32 = 0.$$

Let us move the origin to a point of symmetry, so that the first degree terms will be absent. Put $x = x' + h$, $y = y' + k$, then we have

$$3x'^2 + 6hx' + 3h^2 - 7x'y' - 7hy' - 7kx' - 7hk + 2y'^2 + 4ky'$$
$$+ 2k^2 + 9x' + 9h + 2y' + 2k - 32 = 0.$$

Now, choosing h and k so that the coefficients of x' and y' are zero, we have

$$6h - 7k + 9 = 0,$$
$$-7h + 4k + 2 = 0,$$

so that $h = 2$ and $k = 3$. The equation now becomes

$$3x'^2 - 7x'y' + 2y'^2 - 20 = 0.$$

Now let us rotate the axes to positions of symmetry, so that the product term may be absent. Let

$$x' = x'' \cos \alpha - y'' \sin \alpha, \; y' = x'' \sin \alpha + y'' \cos \alpha,$$

then we have

$$3x''^2 \cos^2 \alpha - 6x''y'' \sin \alpha \cos \alpha + 3y''^2 \sin^2 \alpha$$

$$-7x''^2 \sin \alpha \cos \alpha - 7x''y'' (\cos^2 \alpha - \sin^2 \alpha) + 7y''^2 \sin \alpha \cos \alpha$$

$$+2x''^2 \sin^2 \alpha + 4x''y'' \sin \alpha \cos \alpha + 2y'' \cos^2 \alpha - 20 = 0.$$

Choosing α so that the coefficient of $x''y''$ is zero, we have

$$-2 \sin \alpha \cos \alpha = 7 (\cos^2 \alpha - \sin^2 \alpha)$$

or

$$\tan 2\alpha = -7.$$

From this we have $\cos 2\alpha = \frac{1}{10}\sqrt{2}$, $\sin^2\alpha = \frac{10 - \sqrt{2}}{20}$, $\cos^2 \alpha = \frac{10 + \sqrt{2}}{20}$; $\sin \alpha \cos \alpha = -\frac{7}{20}\sqrt{2}$, and the equation becomes

$$\frac{5}{2}(1 + \sqrt{2})\, x''^2 + \frac{5}{2}(1 - \sqrt{2})\, y''^2 = 20.$$

The student should make a graph of the original equation and draw the two new sets of axes. Notice that the equation may be written

$$(x - 2y + 4)(3x - y - 3) = 20,$$

and draw the asymptotes first. The point of symmetry is the intersection of the asymptotes, and the axes of symmetry are the bisectors of the angles between the asymptotes.

EXERCISES XXIX

1. Transform to polar coordinates:

(a) $y^2 = 4p(x + p)$,
(b) $x^2 - y^2 = 16$,
(c) $x^2 + y^2 + 6x = 0$.

2. Transform to rectangular coordinates:

(a) $\rho = 8 \cos \theta$; (c) $\rho = \dfrac{5}{\cos\left(\theta - \dfrac{\pi}{3}\right)}$;

(b) $\rho = \dfrac{8}{1 - \cos \theta}$; (d) $\rho \cos \theta = 4$.

3. Given $x^2 + y^2 - 6x + 2y + 1 = 0$, move the origin to a point of symmetry and plot the curve.

4. Rotate the axes through the angle $\dfrac{\pi}{4}$ for the curve $x^2 - y^2 = 0$.

5. Show that the equation $x^2 + y^2 = r^2$ is not changed by rotation.

6. In polar coordinates, the transformation $\theta = \theta' + \alpha$ is equivalent to a rotation of the initial line through the angle α. Verify this for the equation $\rho = 10 \sin \theta$, by making the given transformation and then transforming both new and old equations to rectangular coordinates.

7. Given the equation

$$xy - 6x + 8y + 2 = 0,$$

transform by translation and rotation to the form

$$x^2 - ky^2 = c$$

where k and c are constants.

8. Given the equation

$$x^2 + 3x - 3y + 6 = 0,$$

transform by rotation and translation to the form $y^2 = kx$.

Art. 33. The Conic Sections

A very important class of curves is the class called **conic sections.** This name arises from the fact that any curve of this class may be regarded as the curve in which a cone is intersected by a plane, and that any plane cutting a right circular cone cuts it in a curve of this class.

The characteristic property of the conic sections may be shown geometrically by reference to Figure 43. The cone with vertex at V is cut by the plane MN in the curve EHP. A sphere touching the plane MN at F is in contact with the cone at all points of the circle CSD. The axis of the cone is perpendicular to the plane of the circle at its center, O, and cuts the plane MN at B. The plane of the circle, MK,

meets MN in the line QM. The line OA is perpendicular to QM, and, BF is perpendicular to QM at A.

Now take any point P of the curve, draw PF, and draw the element of the cone PV, cutting the plane MK at S. Draw PR parallel to BO, and PQ parallel to BA. Triangle PRS is similar to VOS or VOC, and

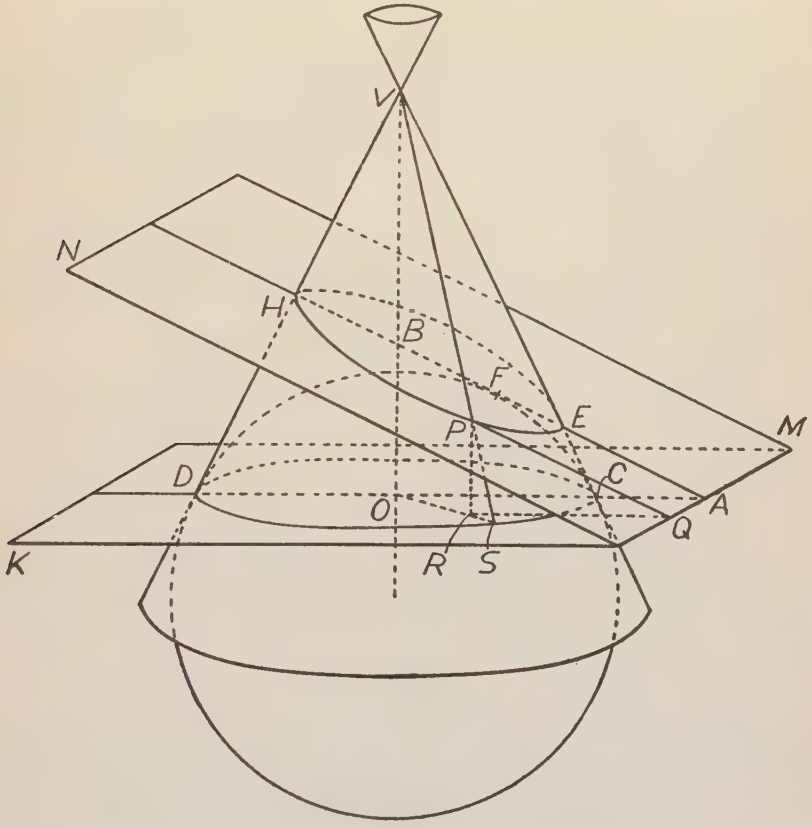


Fig. 43

triangle PQR is similar to BAO. Since PF and PS are tangents to the sphere from P they are equal, therefore

$$\frac{PF}{PQ} = \frac{PS}{PQ} = \frac{PS}{PR} \cdot \frac{PR}{PQ} = \frac{VC}{VO} \cdot \frac{BO}{BA}.$$

Now this last expression does not depend on the position of P on the curve, and therefore the ratio of PF to PQ has the same value for

every point P of the curve. If α is the angle that an element of the cone makes with the axis of the cone, and ϕ is the angle at which the cutting plane MN intersects this axis, then this constant ratio is

$$\frac{PF}{PQ} = \frac{VC}{VO} \cdot \frac{BO}{BA} = \frac{\cos \phi}{\cos \alpha}.$$

This ratio is called the **eccentricity** of the curve, usually denoted by e, the point F is called a **focus,** and the line QM is called a **directrix.** We arrive, then, at the following definition:

A conic section is the locus of a point in a plane whose distance from a fixed point in the plane bears a fixed ratio to its distance from a fixed line in the plane.

To write the rectangular equation of a conic section in conventional form we take the directrix as the y-axis and take the x-axis through the focus, F (Fig. 44), and let the distance from focus to directrix be 2p. Taking $\frac{PF}{PQ} = e$, a constant, we have

$$\sqrt{(x - 2p)^2 + y^2} = ex,$$

which reduces to

$$(1 - e^2)x^2 + y^2 - 4px + 4p^2 = 0. \tag{1}$$

Again, to write the equation in polar coordinates, we take the pole at F and take the initial line through F perpendicular to the directrix. We have, then,

$$\rho = e(2p + \rho \cos \theta)$$

which reduces to

$$\rho = \frac{2ep}{1 - e \cos \theta}. \tag{2}$$

To discuss the properties of the conic sections it is convenient to consider the several different types of curves of this class.

1. **The circle.** If in Figure 43 the cutting plane MN is allowed to move into a position perpendicular to the axis VB at B, the curve EHP becomes a circle, the eccentricity becomes zero, but the line QM recedes to a distance indefinitely great. For the circle, then the directrix is "at infinity," and the focus is at the center. This is a special case, then, in which the equation (1) above does not apply. Since all

points of the circle are equally distant from the center, we have, taking the center at (h, k) and the radius as r,

$$(x - h)^2 + (y - k)^2 = r^2, \tag{3}$$

and when h and k are zero this is in its simplest form.

If we put $e = 0$ in equation (1) we have

$$(x - 2p)^2 + y^2 = 0,$$

and the locus of this equation has the single real point (2p, 0). This corresponds to the case when the plane MN (Fig. 43) passes through

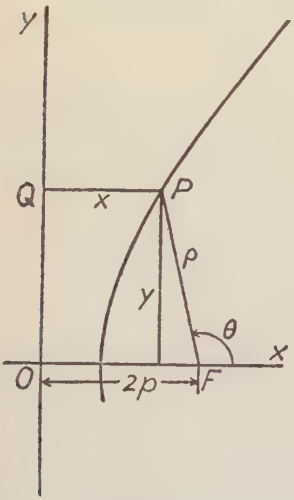


Fig. 44

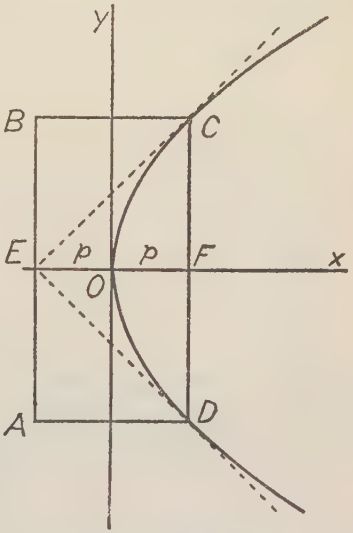


Fig. 45

V perpendicular to VB. The locus is known as a **point circle,** and corresponds to equation (3) with $r = 0$.

2. **The parabola.** In equation (1) put $e = 1$, then we have

$$y^2 - 4px + 4p^2 = 0,$$

or

$$y^2 = 4p(x - p).$$

If we transform by translation to the point (p, 0) the equation takes the form

$$y^2 = 4px, \tag{4}$$

which is the standard equation of a **parabola,** and from it the properties of the parabola may be studied. The curve (Fig. 45) is symmetrical with respect to the x-axis, which is called the **axis** of the parabola, and the locus lies entirely to the right of the y-axis. The rectangle ABCD has dimensions 2p and 4p, and the line AB is the directrix while the line CD, of length 4p, is called the **latus rectum** of the parabola. The point O at which the parabola cuts its axis is called its **vertex.** All parabolas are the same shape, and differ only in size, the size being controlled by the value of p. The quantity p is, in fact, a parameter of a family of parabolas, and may take any value, positive or negative. For negative values of p the locus must lie to the left of the y-axis.

The equation

$$x^2 = 4py, \tag{4'}$$

by comparison with (4), is seen to represent a parabola with vertex at the origin and with axis on the y-axis, and the curve may lie either above or below the x-axis according as p is positive or negative.

The equations

$$(y - k)^2 = 4p(x - h), \tag{4''}$$

$$(x - h)^2 = 4p(y - k), \tag{4'''}$$

obviously represent parabolas with vertex at (h, k) and with axes horizontal and vertical, respectively.

If e = 1 in equation (2) we have

$$\rho = \frac{2p}{1 - \cos\theta} = p \csc^2 \frac{\theta}{2}, \tag{5}$$

and these are two common forms of the **polar equation** of a **parabola.** The equations

$$\rho = \frac{2p}{1 + \cos\theta} = p \sec^2 \frac{\theta}{2} \tag{5'}$$

represent a parabola lying to the left of its vertex, or rotated through the angle π, since $\cos\theta = -\cos(\theta + \pi)$ and $\csc^2 \frac{\theta}{2} = \sec^2 \frac{(\theta + \pi)}{2}$.

3. **The ellipse.** If e is not unity, then equation (1) may be written

$$\frac{\left(x - \frac{2p}{1 - e^2}\right)^2}{\frac{4e^2p^2}{(1 - e^2)^2}} + \frac{y^2}{\frac{4e^2p^2}{1 - e^2}} = 1. \tag{1'}$$

Now if e is between zero and unity we may write

$$a^2 = \frac{4e^2p^2}{(1 - e^2)^2}, \quad b^2 = \frac{4e^2p^2}{1 - e^2}, \tag{6}$$

and the equation, after shifting the origin to the point $\left(\frac{2p}{1 - e^2}, 0\right)$, takes the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{7}$$

This is the equation of an **ellipse** in its simplest standard form, and it corresponds to the curve in Figure 46 referred to axes OA and OB.

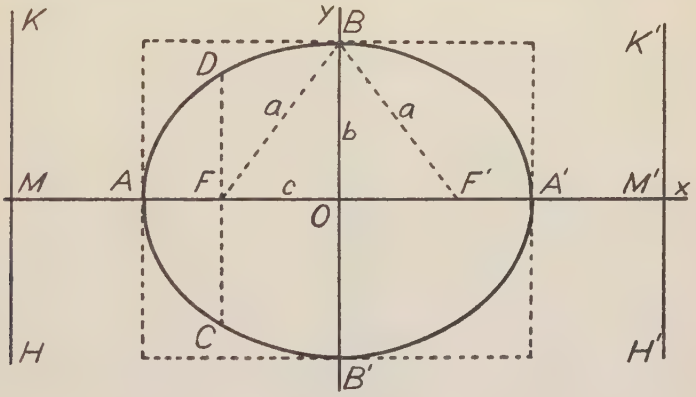


Fig. 46

The curve is symmetrical with respect to both axes and the origin, the x-intercepts are a and $-a$, the y-intercepts are b and $-b$. The line AA′, of length 2a, is the **major axis,** and the line BB′, of length 2b, is the **minor axis.** If HK is the directrix and F is the focus, we have $OM = \frac{2p}{1 - e^2}$, and as B is a point on the curve

$$BF = e \cdot OM = \frac{2ep}{1 - e^2} = a,$$

and since $\frac{b^2}{a^2} = 1 - e^2$ we have

$$e^2 = \frac{a^2 - b^2}{a^2}, \text{ or } e = \frac{c}{a}, \tag{8}$$

where $c = OF$. If we put $x^2 = c^2 = a^2 - b^2$ in equation (7) we have $y = \pm \frac{b^2}{a}$, showing that the double ordinate, CD, through F is of length $\frac{2b^2}{a}$. This double ordinate is called the **latus rectum** of the ellipse. By symmetry we see that this ellipse has another directrix at H′K′ and another focus at F′.

With any given ellipse, with semi-axis a and b, there are associated two circles, one inscribed with radius b, called the **minor auxiliary circle,** the other circumscribed with radius a, called the **major auxiliary circle** (Fig. 47).

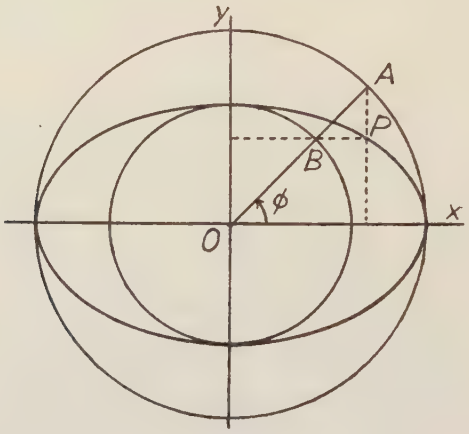


Fig. 47

The equations

$$x = a \cos \phi, \quad y = b \sin \phi, \tag{9}$$

are **parametric equations** of the ellipse, since elimination of ϕ gives equation (7). By consideration of these equations a simple construction to locate a point P of an ellipse is at once evident. Through O (Fig. 47) draw a line making any angle ϕ with Ox, and intersecting major and minor auxiliary circles at A and B respectively. A vertical line through A and a horizontal line through B intersect at the required point.

Equation (2), with e less than unity, is the polar equation of an ellipse with the pole at the left hand focus. If the negative sign in the denominator is changed to positive, the pole becomes the right hand focus.

To obtain the polar equation of an ellipse with pole at the center we may transform equation (7) to polar coordinates, obtaining

$$\rho^2 = \frac{a^2b^2}{a^2 \sin^2 \theta + b^2 \cos^2 \theta}. \tag{10}$$

The equation

$$\frac{(x - h)^2}{a^2} + \frac{(y - k)^2}{b^2} = 1 \tag{10'}$$

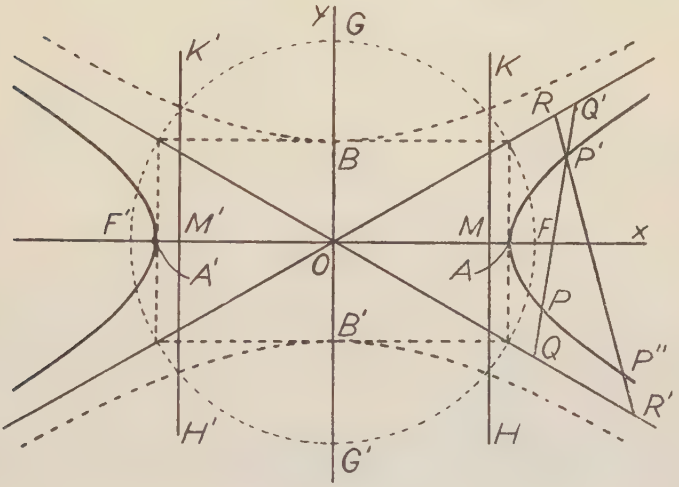


FIG. 48

obviously represents an ellipse with center at (h, k) with semi-axes a and b, with the major axis horizontal if a is greater than b vertical, if a is less than b. The equation becomes that of a circle if a = b.

4. **The hyperbola.** If e is greater than unity, then in equation (1') the quantity $1 - e^2$ is negative and we may write

$$a^2 = \frac{4e^2p^2}{(1 - e^2)^2}, \quad b^2 = -\frac{4e^2p^2}{1 - e^2}. \tag{11}$$

and, shifting the origin as before to $\left(\frac{2p}{1 - e^2}, 0\right)$, the equation becomes

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{12}$$

This is the equation of a hyperbola in standard form. The curve is symmetrical with respect to both axes and with respect to the origin, the x-intercepts are a and $-a$, and for x between these values there are no real values of y (Fig. 48). If HK is the directrix and F is the focus, then $OM = \frac{2p}{e^2 - 1} = \frac{a}{e}$. Now since A is a point on the curve

$$\frac{AF}{AM} = e = \frac{a}{OM'}$$

and, therefore, using composition

$$\frac{AF + a}{AM + OM} = \frac{OF}{a} = e$$

But, since $\frac{b^2}{a^2} = e^2 - 1$, we have

$$e^2 = \frac{a^2 + b^2}{a^2} \text{ or } e = \frac{c}{a'} \tag{13}$$

where $c = OF$ and $2c$ is the diagonal of the rectangle whose dimensions are $2a$ and $2b$. The line AA' is called the **transverse axis** of the hyperbola, and BB' the **conjugate** axis.

If in equation (8) we put $x^2 = c^2 = a^2 + b^2$, we have $y = \pm \frac{b^2}{a}$, showing that the latus rectum of the hyperbola, the double ordinate through F or F′, has a length $\frac{2b^2}{a}$.

Equation (12) may be written

$$\left(\frac{x}{a} - \frac{y}{b}\right)\left(\frac{x}{a} + \frac{y}{b}\right) = 1, \tag{12′}$$

showing by the principles of Art. 31 that the lines $bx - ay = 0$ and $bx + ay = 0$, the diagonals of the rectangle, are **asymptotes to the hyperbola,** and that the hyperbola is the locus of points the product of whose distances from these lines is constant.

$$\frac{bx - ay}{\sqrt{a^2 + b^2}} \cdot \frac{bx + ay}{\sqrt{a^2 + b^2}} = \frac{a^2b^2}{a^2 + b^2}.$$

This last property of the hyperbola provides a convenient method of finding points on the curve, given the asymptotes and one point,

as P. Draw a line through P, cutting the asymptotes at Q and Q′, then take Q′P′ = QP, and P′ is a point on the curve; for it is evident geometrically that the product of distances to the asymptotes is the same from P′ as from P. Similarly, let RR′ through P′, cut the asymptotes at R and R′ then if R′P″ = RP′, the point P″ is on the curve.

Equation (2), with e greater than unity, is the polar equation of a hyperbola with pole at the right hand focus. If the negative sign in the denominator is changed to positive the pole becomes the left hand focus, F′. The polar equation with pole at the center of the hyperbola, corresponding to equation (10) for the ellipse, is

$$\rho^2 = \frac{a^2b^2}{b^2\cos^2\theta - a^2\sin^2\theta}. \tag{14}$$

Convenient parametric equations for the hyperbola are:

$$x = a\sec t, \quad y = b\tan t; \tag{15}$$

or

$$x = a\cosh t, \quad y = b\sinh t. \tag{16}$$

Elimination of t from either of these sets of equations will give equation (12).

By comparison with (12) it will be seen that the equation

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1 \tag{17}$$

represents a hyperbola, shown as dotted curve in Figure 48, having the same asymptotes but having foci at G and G′ on the y-axis. The two curves, (12) and (17), are **conjugate** hyperbolas.

If a = b in (17) or (12), then the rectangle in Figure 48 becomes a square, the asymptotes intersect at right angles, and the curve is called a **rectangular** hyperbola. In this case, by rotation through $\frac{\pi}{4}$, the axes of coordinates become the asymptotes and the equation takes the form

$$xy = \frac{a^2}{2}. \tag{18}$$

It remains to observe that the equations

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = \pm 1. \tag{19}$$

$$(x-h)(y-k) = \pm \frac{a^2}{2}. \tag{20}$$

represent two pairs of conjugate hyperbolas, with center at (h, k).

EXERCISES XXX

1. Write the equations of circles that fulfil conditions as follows:

(a) Center at (2, 5), passing through (5, 1).

(b) Through the points (1, 5) and (−3, 6) with center on the line $3x - 2y = 6$.

(c) Through the points (2, −1), (−5, 2), (3, 3).

2. Write the equations of parabolas as follows:

(a) With vertex at (2, 4), focus at (3, 4)

(b) With axis vertical, passing through the points (0, 1), (2, 3) and (5, −1).

(c) Vertex at origin, directrix the line $x = 3$.

3. Show that the locus of points the sum of whose distances from (c, 0) and (−c, 0) is equal to 2a is an ellipse with the two given points as foci and with intercepts at (a, 0) and (−a, 0).

4. Write the equation of an ellipse with foci at (2, 4) and (8, 4), passing through (5, 0).

5. Write the equation of a hyperbola with asymptotes $3x \pm 4y = 0$ and

(a) With a focus at (0, 10).

(b) With a focus at (5, 0).

(c) Through the point (7, 2).

6. Treat problem 3 after changing "sum" to "difference" and "ellipse" to "hyperbola."

7. Write the equation

$$y^2 = 12x + 6y + 15$$

in the form (4″) and find the vertex, the focus and the directrix of the parabola.

8. Write each of the following equations in a standard form of this article and find all centers, foci and asymptotes, and name each curve.

(a) $x^2 + y^2 + 6x - 8y = 0.$

(b) $x^2 - 4y^2 + 2x + 16y - 11 = 0.$

(c) $x^2 + 4x - 6y + 22 = 0.$

(d) $16x^2 - 96x + 9y^2 + 18y + 9 = 0.$

9. A rectangular hyperbola has the x-axis as one asymptote. It passes through the points (2, 8) and (5, 4). Find the other asymptote.

10. An ellipse has its center at (3, 8) and one focus at (3, 2) and its eccentricity is $\frac{3}{4}$. Write its equation and find the length of its latus rectum.

11. What type of conic section is given by the equation

$$\rho = \frac{16}{3 + 4\cos\theta},$$

and where is its center?

12. Discuss the locus given by

$$x = 2 + 5\cos\phi, \quad y = -3 + 3\sin\phi.$$

13. Write parametric equations for the hyperbola

$$x^2 - 9y^2 = 36.$$

Art. 34. The Equation of the Second Degree

It may be noticed that the rectangular equations of conic sections considered in the last article are all of the second degree in x and y. In the present article we give attention to the conditions under which a second degree equation may represent a conic section of given type, and develop methods of analysis by which to determine, in any case, what locus such an equation does represent.

Any parabola, by convenient choice of axes, may have its equation written

$$y^2 = 4px, \tag{1}$$

and similarly any ellipse may be represented by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{2}$$

and any hyperbola by

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{3}$$

Also, any circle may be represented by (2) by making $a = b = r$.

Let us take as the general equation of the second degree

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0, \tag{4}$$

where the six constants may take any real values including zero, provided that A, B and C are not all zero. Since the equation may be multiplied by any constant, there are, in fact, only five independent parameters, and these may be taken as the ratios that five of the constants bear to the sixth, always taking as the sixth a constant which is not zero.

Now if (4) is to represent a circle, parabola, ellipse or hyperbola it must come from (1), (2) or (3) by transformation of coordinates. Translation does not change the terms of the second degree, and rotation changes second degree terms into other terms of the second degree. We see, therefore, that (4) can come from equation (1) only in case

$$Ax^2 + Bxy + Cy^2$$

is a perfect square, so that for a **parabola** we must have $B^2 - 4AC = 0$. Also, if (4) comes from (2) the second degree terms must be a sum of squares, so that for an **ellipse** we must have $B^2 - 4AC < 0$, and finally, for a **hyperbola** we must be able to factor the second degree terms into two real linear factors, so that we must have $B^2 - 4AC > 0$.

Thus, by a moment's inspection of the second degree terms we are able to assign equation (4) to one of the three types of curves represented respectively by (1), (2) and (3). In giving the methods of treating equation (4) to find the exact character of the curve it represents it is convenient to consider the three cases separately.

1. When $B^2 - 4AC = 0$. If, now, we have $B = 0$, then either $A = 0$ or $C = 0$, and the equation may be reduced to the form (4″) or (4‴) of Art. 33, from which the curve is completely known. If B is not zero, we may rotate the axes so that in the new equation the product term is missing, then proceed as before. Under certain conditions, however, the equation may represent two straight lines, parallel, coincident, or imaginary. We may illustrate as follows: Let the

equation, after rotating, if necessary, to dispose of the product term, take the form

$$y^2 + Dx + Ey + F = 0,$$

then we may write

$$y^2 + Ey + \frac{E^2}{4} = -D\left(x - \frac{E^2 - 4F}{4D}\right)$$

and the equation is in the form of (4″) of Art. 33. But if $D = 0$, we have

$$y = -\frac{E}{2} \pm \frac{1}{2}\sqrt{E^2 - 4F},$$

which represents two horizontal lines, distinct, coincident or imaginary according as $E^2 - 4F$ is positive, zero or negative. If x^2 is present instead of y^2, a similar treatment leads to form (4‴) of Art. 33 or to vertical lines.

Frequently it is convenient to inspect an equation without reduction to standard form. For example, take

$$9x^2 + 12xy + 4y^2 + 30x - 32y + 25 = 0.$$

We may write

$$(3x + 2y)^2 + D(2x - 3y) + E(3x + 2y) + F = 0,$$

and proceed to determine D, E and F so that this equation agrees with the given equation. This gives $D = 12$, $E = 2$, $F = 25$, and proceeding as above, but taking $3x + 2y$ as a new y and $2x - 3y$ as a new x,

$$(3x + 2y)^2 + 2(3x + 2y) + 1 = -12(2x - 3y) - 24$$

or

$$(3x + 2y + 1)^2 = -12(2x - 3y + 2).$$

It is now clear that this is a parabola, and that it may be reduced to standard form by taking as new axes the lines

$$3x + 2y + 1 = 0 \text{ and } 2x - 3y + 2 = 0.$$

To accomplish this we may take

$$3x + 2y + 1 = \sqrt{13}\,y', \qquad 2x - 3y + 2 = \sqrt{13}\,x'$$

and the equation becomes

$$y'^2 = -\frac{12}{\sqrt{13}}x'.$$

In some cases, in solving for D, E and F we may find D = 0, then the equation gives straight lines, their character depending on E and F as before.

2. When $B^2 - 4AC < 0$. Here, if B = 0, we must have A and C of like sign. If A = C, the locus is a circle, and may at once be reduced to standard form. For example

$$x^2 + y^2 - 6x + 4y - F = 0$$

may be written

$$x^2 - 6x + 9 + y^2 + 4y + 4 = F + 13$$

or

$$(x - 3)^2 + (y + 2)^2 = F + 13$$

showing that the locus is a circle of radius $\sqrt{F + 13}$, with center at (3, −2). This circle will be real, a point circle or an imaginary circle according as F + 13 is positive, zero or negative. To illustrate when A and C are not equal, take

$$4x^2 + 9y^2 + 8x - 36y - F = 0.$$

This may be written

$$4(x^2 + 2x + 1) + 9(y^2 - 4y + 4) = F + 40.$$

This is an ellipse, real, a single point or imaginary according as F + 40 is positive, zero or negative. If, in particular, F = −4, we may write

$$\frac{(x + 1)^2}{9} + \frac{(y - 2)^2}{4} = 1,$$

and the locus is an ellipse with center at (−1, 2), with semi-axes 3 and 2.

If B is not zero, we may rotate the axes to obtain a new equation free from the product term. In this case, however, it will usually be more convenient to move the origin to the center first, by translation, then the first degree terms will be absent. The equation now takes the form

$$Ax^2 + Bxy + Cy^2 = K,$$

where A, B and C are unchanged, but K is a new constant. The equation now represents a real ellipse, a point ellipse or an imaginary locus according as K is positive, zero or negative. If the locus is real, it may be reduced to standard form by rotation through the angle α, where $\tan 2\alpha = \dfrac{B}{A - C}$.

3. When $B^2 - 4AC > 0$. In this case, if $B = 0$, A and C must be of opposite signs, and we may take as an example

$$4x^2 - 9y^2 + 8x - 36y - F = 0,$$

and write

$$4(x^2 + 2x + 1) - 9(y^2 + 4y + 4) = F - 32.$$

If $F = 32$ this represents two intersecting lines,

$$2(x + 1) = \pm 3(y + 2).$$

If F is not equal to 32, these lines are the asymptotes of the hyperbola, and the transverse axis is horizontal or vertical according as $F - 32$ is positive or negative. If $F = -4$, we may write

$$\frac{(x + 1)^2}{9} - \frac{(y + 2)^2}{4} = -1,$$

representing a hyperbola with center at $(-1, -2)$, with semi-axes 3 and 2, with foci at $(-1, -2 \pm \sqrt{13})$.

If B is not zero we may rotate to dispose of the product term, then proceed as before. But it is frequently more convenient and useful to find the asymptotes and sketch the curve by the following method of inspection. The second degree terms may be factored, and we may take as an illustration,

$$(4x + 5y)(2x - 7y) + 8x + 29y - 24 = 0.$$

Let us write

$$(4x + 5y + p)(2x - 7y + q) = c,$$

and proceed to determine p, q and c so that this equation may agree with the given equation. This requires $p = -2$, $q = 3$, $c = 18$, so that the asymptotes are

$$4x + 5y - 2 = 0, \quad 2x - 7y + 3 = 0,$$

and the product of distances from these lines to any point of the curve is

$$\frac{4x + 5y - 2}{\sqrt{41}} \cdot \frac{2x - 7y + 3}{\sqrt{53}} = \frac{18}{\sqrt{41}\sqrt{53}}.$$

If in applying this method we should find $c = 0$, then the locus consists of the two lines.

EXERCISES XXXI

1. Reduce each of the following to standard form and give values of standard parameters.

(a) $x^2 + y^2 - 8y - 9 = 0.$
(b) $x^2 + 6x - 8y - 7 = 0.$
(c) $x^2 - 4y^2 - 24y + 24 = 0.$
(d) $4x^2 + y^2 + 8x - 2y - 31 = 0.$

2. Show that each of the following equations represents two straight lines.

(a) $16x^2 + 8xy + y^2 + 12x + 3y + 2 = 0.$
(b) $x^2 - xy - 2y^2 - x - 4y - 2 = 0.$

3. Show that each of the following loci are imaginary.

(a) $x^2 + y^2 - 4x + 8y + 26 = 0.$
(b) $x^2 - xy + y^2 + 3x - 3y + 5 = 0.$
(c) $x^2 - 2xy + y^2 + x - y + 1 = 0.$

4. Reduce to standard form by transformation of coordinates, draw the curve and draw both new and old axes.

(a) $x^2 - 2xy + y^2 + 2x - y - 1 = 0.$
(b) $5x^2 + 2xy + 5y^2 - 12x - 12y = 0.$

5. Determine the nature of the following loci.

(a) $y^2 - xy - 5x + 5y = 0.$
(b) $x^2 - 10xy + y^2 + x + y + 1 = 0.$
(c) $x^2 - 3xy + y^2 + 10x - 10y + 21 = 0.$
(d) $9x^2 - 24xy + 16y^2 - 18x - 101y + 19 = 0.$

Art. 35. Certain Families of Curves and Lines

The general equation of the second degree represents a five-parameter family of curves, but a choice of axes involves fixing three parameters, as two coordinates of the new origin and the angle of rotation, so that only two parameters pertain to the intrinsic properties of the conic

section itself. These parameters control shape and size of the curve. For the parabola and circle the shape is fixed, the eccentricities being respectively unity and zero, so that for these curves a single parameter determines the only variable property of the curve, namely its size, the parameter for the parabola being its latus rectum and for the circle its radius.

The ellipse and hyperbola possess the two variable properties, shape and size, and these are determined by two parameters, the quantities p and e of the general equation, or a and b of the simple standard equations. The equation

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1. \tag{1}$$

represents a family of **confocal conics,** if c is regarded as fixed while a is regarded as the parameter of the family. The foci are at (c, 0) and (−c, 0), and when a is greater than c the curve is an ellipse, while for a less than c the curve is a hyperbola.

If the equations

$$f_1(x, y) = 0, \qquad f_2(x, y) = 0,$$

represent conic sections, then the equation

$$mf_1(x, y) + nf_2(x, y) = 0, \tag{2}$$

represents another conic section through their common points, however m and n are chosen. Only the ratio of m to n is significant, since any factor may be divided out. This is, therefore, a one-parameter family of conic sections, and one other condition will completely determine the equation (2) when the original curves are given. The equations being of the second degree, there are four common points, which may or may not be finite and real. A special case worthy of notice is when the two given equations represent circles. In this case two of the common points are at infinity, the other two may be real or imaginary. Let the circles be

$$x^2 + y^2 + a_1x + b_1y + c_1 = 0,$$

$$x^2 + y^2 + a_2x + b_2y + c_2 = 0.$$

Then the equation (2) becomes

$$m(x^2 + y^2 + a_1x + b_1y + c_1) + n(x^2 + y^2 + a_2x + b_2y + c_2) = 0,$$

which represents all circles through the points common to the given circles, and in the particular case when $m = -n$, it becomes two lines, one at infinity, the other through the two finite points common to the circles, always a real line.

Figure 49 represents the case when the finite common points are real, Figure 50 the case when they are imaginary. The vertical line, in each figure is the finite part of the locus when $m = -n$. This line is known as the **radical axis** of the given circles, or, if the intersections are real, as the **common chord.** In either figure any two of the circles may be taken as the original ones, to determine the system. In figure 50 the circle through Q_1 and Q_2 is not one of the system, but intersects each circle of the system orthogonally (perpendicularly). If a system of the kind in Figure 50 is superimposed on a system of the kind in

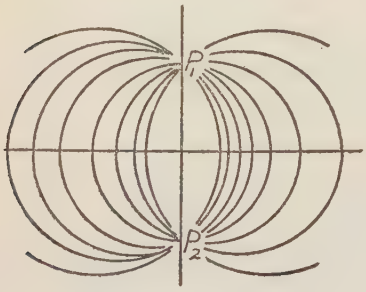


FIG. 49

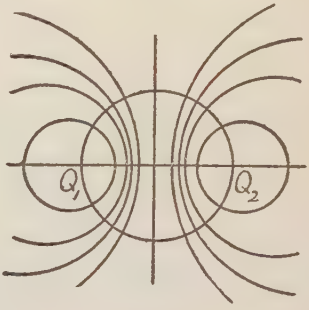


FIG. 50

Figure 49, in such manner that Q_1 and Q_2 coincide with P_1 and P_2, every intersection is at right angles, and the two are said to be orthogonal systems. The orthogonality may be shown most easily by consideration of another property of the radical axis. If (x, y) is any point in the plane, the square of the length of a tangent from this point to the circle with center at (h, k) and radius r is (Fig. 51)

$$(x - h)^2 + (y - k)^2 - r^2,$$

and similarly, the square root of

$$(x - h_1)^2 + (y - k_1)^2 - r_1^2$$

is the length of tangent to another circle with center at (h_1, k_1) and radius r_1. The locus of points (x, y), from which tangents to the two circles are of equal length, found by equating the two lengths written above, is clearly the radical axis of the two circles. From this property

the orthogonal relations mentioned above may easily be deduced, for if from a point C on the radical axis, of one family tangents are drawn to the circles of this family, the points of tangency lie on a circle of the other family with center at C and with radius equal to the common tangent length (Fig. 52). Evidently every circle of this latter family must pass through the common points P and Q.

A linear equation in x and y and an equation of the second degree in the same variables determine two sets of values of x and y, which are the coordinates of the two points of intersection of the line and the curve represented. The line is said to be **tangent** to the second degree curve if the two points of intersection coincide. For example, consider the intersections of the line $y = mx + n$ with the hyperbola

$$x^2 + xy - 2x + 3 = 0.$$

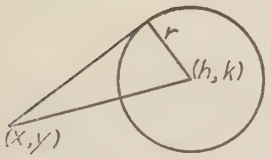


Fig. 51

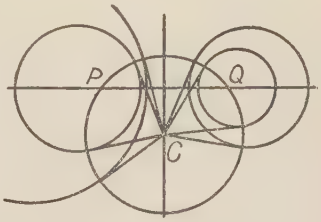


Fig. 52

Eliminating y we have

$$(1 + m)x^2 + (n - 2)x + 3 = 0.$$

If the line is a tangent, the two values of x must coincide, hence

$$(n - 2)^2 = 12(1 + m),$$

$$n = 2 \pm 2\sqrt{3(1+m)}.$$

The equation

$$y = mx + 2 \pm 2\sqrt{3(1 + m)}$$

represents two tangents to the given hyperbola, with slope m. The slope is arbitrary, but to give real tangents must be such as to make the quantity $1 + m$ positive. Any other condition on the line, consistent with the condition of tangency, may be used to determine the value of m. For instance, if it is desired to find a tangent through

(2, 3), we may put $x = 2$ and $y = 3$ in the equation of the tangent and solve for m. We obtain $m = 2 \pm \frac{3}{2}\sqrt{3}$, and the lines through (2, 3) with these slopes are tangents as desired. They are

$$y - 3 = (2 + \tfrac{3}{2}\sqrt{3})(x - 2) \text{ and } y - 3 = (2 - \tfrac{3}{2}\sqrt{3})(x - 2).$$

A **normal** to a curve is a line perpendicular to the tangent at its point of intersection with the curve. Given the equation of a tangent with arbitrary slope, m, in the form,

$$y = mx + f(m),$$

where $f(m)$ is the value of the y-intercept determined as n was determined in the illustration above, we may find, in terms of m, the coordinates of the points of tangency. Denote the coordinates by $f_1(m)$, $f_2(m)$, then the normal at this point is the line through the point with slope $-1/m$. The equation

$$y - f_2(m) = -\frac{1}{m}[x - f_1(m)]$$

is a normal for all values of m. It is desirable, however, to replace m by $-1/m'$, so that the equation of the normal will be written in terms of its own slope m′. Any other condition that a normal may fulfil may be used to determine its slope, as was illustrated above in the case of the tangent.

To illustrate again, take the standard equation of an ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

or

$$b^2x^2 + a^2y^2 - a^2b^2 = 0.$$

Eliminating y between this and $y = mx + n$, we have

$$(m^2a^2 + b^2)x^2 + 2a^2mnx + a^2(n^2 - b^2) = 0.$$

From this we obtain

$$x = -\frac{a^2mn + a\sqrt{a^2m^2n^2 - (m^2a^2 + b^2)(n^2 - b^2)}}{m^2a^2 + b^2}.$$

The line will be tangent to the ellipse if the quantity under the radical is zero and this occurs if

$$n = \pm \sqrt{a^2m^2 + b^2},$$

so that the line, when tangent to the ellipse, must have the form

$$y = mx \pm \sqrt{a^2m^2 + b^2}.$$

Now this touches the ellipse at

$$x = \mp \frac{a^2m}{\sqrt{a^2m^2 + b^2}}, \qquad y = \pm \frac{b^2}{\sqrt{a^2m^2 + b^2}},$$

and the normal through this point is

$$y \mp \frac{b^2}{\sqrt{a^2m^2 + b^2}} = -\frac{1}{m}\left[x \pm \frac{a^2m}{\sqrt{a^2m^2 + b^2}}\right],$$

or

$$y = -\frac{1}{m}x \pm \frac{b^2 - a^2}{\sqrt{a^2m^2 + b^2}},$$

and replacing the slope $-1/m$ of the normal by m' we have for the equation of the normal

$$y = m'x \pm \frac{m'(a^2 - b^2)}{\sqrt{a^2 + b^2m'^2}}.$$

The student may verify the equations of tangents and normals to the standard curves set opposite them in the following table:

Curve	Tangent	Normal
$y^2 = 4px,$	$y = mx + \frac{p}{m},$	$y = mx - 2pm - pm^3;$
$x^2 = 4py,$	$y = mx - pm^2,$	$y = mx + 2p + \frac{p}{m^2};$
$x^2 + y^2 = r^2,$	$y = mx \pm r\sqrt{m^2 + 1},$	$y = mx;$
$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$	$y = mx \pm \sqrt{a^2m^2 + b^2},$	$y = mx \pm \frac{m(a^2 - b^2)}{\sqrt{a^2 + m^2b^2}};$
$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$	$y = mx \pm \sqrt{a^2m^2 - b^2},$	$y = mx \pm \frac{m(a^2 + b_2)}{\sqrt{a^2 - m^2b^2}};$
$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1,$	$y = mx \pm \sqrt{b^2 - a^2m^2},$	$y = mx \mp \frac{m(b^2 + a^2)}{\sqrt{m^2b^2 - a^2}};$
$xy = c,$	$y = mx \pm 2\sqrt{-cm},$	$y = mx \pm \left(\frac{1 - m^2}{m}\right)\sqrt{mc}.$

It should be noticed that the quantity m which enters the equations of the normals of this table is the slope of the normal in question, and is independent of the slope of the tangent. Equations of tangents and normals for secondary standard forms may be written at once. For example, for the ellipse

$$\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1,$$

the tangent and normal are found by putting $x = x' - h$, $y = y' - k$ in the equations set opposite the equation of the ellipse in the table.

These equations in terms of m each represent a family of lines, m being the parameter of the family, every member of the family being tangent, or normal, to the curve. If it is desired to select a member of the family that shall fulfill any second condition, this second condition must furnish an equation by which we may find the value of m. The equations are useful in finding tangents and normals through given points, the degree of the equation in m when particular values are taken for x and y indicating the number of solutions. For example, the equation of a tangent to the parabola gives a quadratic equation for m, showing that there are two tangents through a point, while the normal equation gives a third degree equation for m, showing that three normals may be drawn through a given point, and that at least one of these is real.

EXERCISES XXXII

1. Write the equation of a tangent to the parabola $y^2 = 8x - 4$ with the slope 4. Write the equation of a normal to the same curve, perpendicular to the required tangent.

2. Find the equations of tangents to the circle $(x - 2)^2 + y^2 = 25$ through the point $(-4, 2)$.

3. Find the radical axis of the circles $x^2 + y^2 - 6x + 8y = 0$ and $x^2 + y^2 + 2x - 10y = 10$. Find length of tangents to these circles from the point on the radical axis where $x = 8$.

4. Find the equation of a second degree curve through the points of intersection of $x^2 + 3xy = 2$ and $y^2 + 4x^2 = 5$, and also through the point $(2, 8)$.

5. Find the equation of a parabola through the common points of the two given curves in 4.

6. Find the point from which tangents to the three circles

$$x^2 + y^2 = 25, \quad x^2 + y^2 - 4x + 2y = 4, \quad x^2 + y^2 + 16x + 48 = 0$$

are all of equal length. This point is known as the radical center of the three circles.

7. Find the equations of tangents to the ellipse $9x^2 + 4y^2 = 36$ through the point $(5, -1)$.

9. Write the equation of a family of confocal conics with foci at $(0, 4)$ and $(0, -4)$. Assign two values to the parameter; one giving an ellipse, the other an hyperbola. Find a point of intersection of these two curves, and prove that a tangent to one of these curves at this point is a normal to the other.

10. Show that the two tangents drawn to a parabola from any point in its directrix are perpendicular to each other.

CHAPTER VIII

CERTAIN USES OF DERIVATIVES

Art. 36. Tangents and Normals to Curves

1. **Rectangular Coordinates:** In Article 8 we saw that if $y = f(x)$ is represented by a graph, the derivative $\frac{dy}{dx}$ or $f'(x)$ at any value of x is so related to the tangent line to the graph at the corresponding point that the value of the derivative and the direction of the tangent line are both determined when either is known. In particular, if x and y are measured in the same scale, as in analytic geometry, the value of the derivative is equal to the slope of the line.

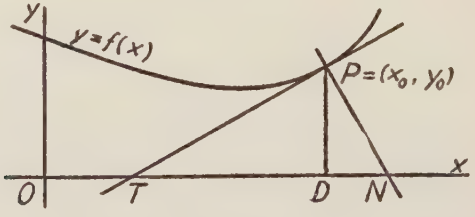


Fig. 53

In the figure (Fig. 53) let the curve be a portion of the graph of $y = f(x)$, in the neighborhood of the point P whose coordinates are x_o, y_o. The derivative at $x = x_o$, that is $f'(x_o)$, is equal to $\frac{DP}{TD}$, the tangent of the angle DTP, the slope of the tangent line TP.

The tangent line passes through the point (x_o, y_o), and has the slope $f'(x_o)$, therefore, using the point-slope equation of Art. 31, the equation of the line is,

$$y - f(x_o) = f'(x_o)(x - x_o), \tag{1}$$

and the equation of the normal line, NP, perpendicular to the tangent at P, is

$$y - f(x_o) = -\frac{1}{f'(x_o)}(x - x_o). \tag{2}$$

The derivative at a point (x_o, y_o) is variously indicated by the notations

$$f'(x_o), \left[\frac{dy}{dx}\right]_{(x_o, y_o)}, \left[\frac{dy}{dx}\right]_o, \frac{dy_o}{dx_o}.$$

The first of these is satisfactory if the equation is given in the explicit form $y = f(x)$, but if the relation between x and y is given in implicit form one of the other three forms is preferable, since they may indicate functions of both x_o and y_o. In this case the equations of tangent and normal are most useful in the forms

$$y - y_o = \frac{dy_o}{dx_o}(x - x_o), \tag{3}$$

$$y - y_o = -\frac{dx_o}{dy_o}(x - x_o). \tag{4}$$

In the use of these equations it must not be forgotten that the derivative is to be evaluated at $x = x_o$ and $y = y_o$, and that these lines are tangent and normal to the curve only when x_o and y_o satisfy the equation of the curve.

If the relation between x and y is stated in parametric form, as

$$x = f_1(t), \quad y = f_2(t),$$

then, if t_o is the value of t for which $x = x_o$ and $y = y_o$, the tangent and normal are given by

$$y - f_2(t_o) = \frac{f_2'(t_o)}{f_1'(t_o)}[x - f_1(t_o)], \tag{5}$$

$$y - f_2(t_o) = -\frac{f_1'(t_o)}{f_2'(t_o)}[x - f_1(t_o)]. \tag{6}$$

Given the values $y_o = DP$, and the slope $\frac{dy_o}{dx_o} = \frac{DP}{TD}$, the values of four other quantities may at once be computed by reference to the figure. They are: PT, the **tangent length**; NP, the **normal length**; TD, the **subtangent**; and DN, the **subnormal.**

2. **Polar Coordinates.** Any relation between ρ and θ, which may be expressed by an equation involving these variables, as $f(\rho, \theta) = 0$, determines a locus for the point P, whose coordinates are (ρ, θ). The direction of the curve at the point P (Fig. 54) is that of the tangent

line TP, and is known when the angle α or one of its functions is known. Since $\alpha = \theta + \psi$, it will be sufficient to determine the tangent of ψ, the angle between the tangent line and the radius vector at P. To find an expression for $\tan \psi$, consider another point, R, whose coordinates are $\rho + \Delta\rho$ and $\theta + \Delta\theta$ and consider the secant line PR, and a circular arc with center at O and radius ρ, cutting OR in Q. Let OS cut the secant line at S and be parallel to the chord PQ. Now $\rho\Delta\theta =$ arc PQ and $\Delta\rho =$ QR, therefore

$$\rho\frac{\Delta\theta}{\Delta\rho} = \frac{\text{arc PQ}}{\text{QR}} = \frac{\text{arc PQ}}{\text{chord PQ}} \cdot \frac{\text{chord PQ}}{\text{QR}} = \frac{\text{arc PQ}}{\text{chord PQ}} \cdot \frac{\text{OS}}{\text{OR}}.$$

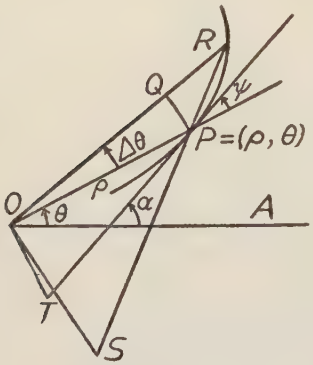


Fig. 54

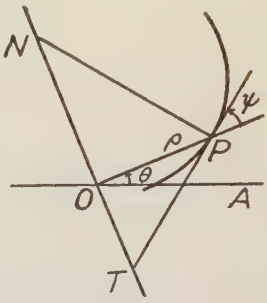


Fig. 55

As $\Delta\rho$ and $\Delta\theta$ approach zero, the triangle ORS approaches the triangle OPT, the angle TOP being a right angle. As we saw in Art. 11, the ratio of the arc to the chord approaches unity, and therefore

$$\rho\frac{d\theta}{d\rho} = \frac{\text{OT}}{\text{OP}} = \tan\psi. \tag{7}$$

For the slope of the tangent line we have

$$\tan\alpha = \tan(\theta + \psi) = \frac{\tan\theta + \tan\psi}{1 - \tan\theta\tan\psi},$$

therefore,

$$\tan\alpha = \frac{\tan\theta + \rho\dfrac{d\theta}{d\rho}}{1 - \rho\tan\theta\dfrac{d\theta}{d\rho}}. \tag{8}$$

This result may be reached by a transformation from rectangular coordinates, if we take the x-axis as OA and the y-axis perpendicular to OA at O, so that the rectangular coordinates of P are

$$x = \rho \cos \theta, \quad y = \rho \sin \theta.$$

We have now

$$\tan \alpha = \frac{dy}{dx} = \frac{\dfrac{dy}{d\rho}}{\dfrac{dx}{d\rho}} = \frac{\sin \theta + \rho \cos \theta \dfrac{d\theta}{d\rho}}{\cos \theta - \rho \sin \theta \dfrac{d\theta}{d\rho}},$$

and dividing both numerator and denominator by cos θ we reach the previous result.

Four quantities with standard designations appear in the figure (Fig. 55), the line TN being drawn perpendicular to the radius vector OP. These quantities may readily be expressed in terms of ρ and tan ψ, by observing similarity of triangles. They are PT, the **polar tangent length**; OT, the **polar subtangent**; ON, the **polar subnormal**; and NP, the **polar normal length.**

3. **Angle of intersection of two curves.** If two curves intersect at a point P, the angle, ϕ at which they meet is the angle between the tangents to the two curves at the common point. If the two tangents make angles α_1 and α_2 with the x-axis, and angles ψ_1 and ψ_2 with the radius vector drawn to the common point, then $\phi = \alpha_1 - \alpha_2 = \psi_1 - \psi_2$. If the curves are given in rectangular coordinates by explicit, implicit or parametric equations, it will be convenient to use

$$\tan \phi = \tan (\alpha_1 - \alpha_2) = \frac{\tan \alpha_1 - \tan \alpha_2}{1 + \tan \alpha_1 \tan \alpha_2},$$

where tan α_1 is the value of $\dfrac{dy}{dx}$ at P computed from the equation of one curve and tan α_2 the corresponding value computed from the other curve. If the curves are given by equations in polar coordinates, tan ψ_1 and tan ψ_2 may be found and tan ϕ computed from them.

EXERCISES XXXIII.

1. Find the equations of tangent and normal lines to the following curves at the point indicated. Also find in each case the lengths of tangent and normal and the subtangent and subnormal.

(a) $y = x^2 + 3x - 2$, at $x = 1$;

(b) $y = \sin x$, at $x = \frac{\pi}{3}$;

(c) $y^2 + 9x^2 = 25$, at $(-1, 4)$;

(d) $x^2 = e^y$, at $(1, 0)$;

(e) $x^2 + xy - 2y = 5$, at $(1, -4)$;

(f) $\begin{cases} x = t^2, \\ y = \sin \pi t, \text{ at } t = \frac{1}{2}; \end{cases}$

(g) $\begin{cases} x = 1 - 2t, \\ y = t^2 - 4t, \text{ at } t = 1; \end{cases}$

(h) $\begin{cases} x = 4t^2, \\ y = t^3, \text{ at } t = -2. \end{cases}$

2. Find the polar tangent length, polar normal length, polar subtangent and polar subnormal for the following curves at the points indicated:

(a) $\rho = a \sec^2 \frac{\theta}{2}$ at $\theta = \frac{\pi}{2}$; (parabola)

(b) $\rho = a \sin \theta$, at $\theta = \frac{\pi}{3}$; (circle)

(c) $\rho^2 = a^2 \cos 2\theta$, at $\left(-\frac{a}{\sqrt{2}}, \frac{\pi}{6}\right)$. (lemniscate)

3. Show that the subnormal to the parabola $y^2 = 4px$ is $2p$ for all points on the curve, and that the subtangent for any point is bisected at the origin. From these facts prove that the focus of the parabola is equidistant from the point of tangency and the points where tangent and normal meet the x-axis.

4. Show that if (x_1, y_1) and (x_2, y_2) are points on the circle $x^2 + y^2 = r^2$, then $x_1x + y_1y = r^2$ and $x_2x + y_2y = r^2$ are equations of tangents at these points. If these tangents intersect at (x_o, y_o), show that the line $x_ox + y_oy = r^2$ cuts the circle at (x_1, y_1) and (x_2, y_2). Draw a figure, putting $r = 5$, $x_1 = 3$, $y_1 = 4$, $x_2 = 4$, $y_2 = -3$.

5. Show that the hyperbola $x^2 - y^2 = 5$ and the ellipse $\frac{x^2}{18} + \frac{y^2}{8} = 1$ intersect at right angles.

6. Show that the two spirals $\rho = a\theta$ and $\rho = \frac{a}{\theta}$ intersect at right angles, and that the former has a constant polar subnormal and the latter a constant polar subtangent.

7. Let s, v and a denote respectively distance, velocity and acceleration of a moving object. Using the relation

$$a = \frac{dv}{dt} = \frac{dv}{ds}\cdot\frac{ds}{dt} = v\frac{dv}{ds},$$

show that, if v is given as a function of s, the subnormal to the graph of this function at any value of s is equal to the acceleration at that value of s.

8. Show by direct differentiation, and also by use of examples 7 and 3 that if an object moves so that $v^2 = 2as$, the acceleration is constantly equal to a.

9. Find the angle of intersection of the two curves:

$$y = e^x, \quad y = e^{2x}.$$

10. Find the equations of tangent and normal to the ellipse $x = a\cos\phi$, $y = b\sin\phi$ at the point where $\phi = \phi_0$.

11. Show that the normal at any point of an ellipse bisects the angle between the lines drawn from this point to the foci.

12. Find the area bounded by the coordinate axes and a tangent to the hyperbola $xy = 18$. Show that the area is independent of the point of contact.

13. Find the subtangent and subnormal for the hyperbola $x = a\sec t$, $y = b\tan t$, at the point where $t = t_0$.

14. Show that if the two tangents are drawn from a point in the directrix to a parabola, the line joining the points of contact passes through the focus.

Art. 37. Use of Graphs of Derived Functions

1. **Curve tracing:** We have considered in Arts. 3 and 4, and in Chapter VII, methods of constructing graphs of functions of various types. The construction of such graphs may be greatly facilitated by the use of derived functions and their graphs.

Suppose a function, $y = f(x)$, to be represented by a curve or graph, and consider the derived function $y' = f'(x)$, also represented by a graph. Since the value of y' at any x determines the slope of the tangent to the original curve, we see that if $y' = 0$, at $x = x_0$, then the original curve has a horizontal tangent at x_0; that if y' is positive at x_0, y is an increasing function at x_0, i.e., its graph rises as x increases through x_0; that if y' is negative at x_0, y is a decreasing function at x_0. Further, if y' is increasing at x_0, the slope of a tangent to the y curve is increasing, so that

the y curve is concave upward at x_o; and similarly, if y' is decreasing at x_o, the y curve is concave downward at x_o. If y' is increasing as x approaches x_o from the left, but decreasing as x moves to the right from x_o, then the graph of $y = f(x)$ is concave upward to the left of x_o and concave downward to the right of x_o. A point x_o, at which the concavity changes from upward to downward, or the reverse, is a **point of inflexion.**

	$y' = f'(x)$	$y = f(x)$
1.	Zero	Horizontal tangent
2.	Positive	Increasing
3.	Negative	Decreasing
4.	Increasing	Concave upward
5.	Decreasing	Concave downward
6.	Slope changing sign	Point of inflexion.
7.	Tangent to x-axis, concave upward.	Horizontal tangent, rising point of inflexion.
8.	Tangent to x-axis, concave downward.	Horizontal tangent, falling, point of inflexion.
9.	Tangent to x-axis, rising point of inflexion	Horizontal tangent, concave upward.
10.	Tangent to x-axis falling point of inflexion	Horizontal tangent, concave downward.

The table presented here gives in condensed form the simple relations that hold between a function and its derived function, and between their graphs. In entries 7 and 9 the word "rising," used in connection with "horizontal tangent," and "point of inflexion," is to indicate that the graph is rising both to the left and right of the point of inflexion, i.e., the curve crosses its horizontal tangent at that point, going from below to above the tangent. A similar interpretation is intended for "falling" in entries 8 and 10. Entry 7 is deduced from 1, 2 and 6; entry 8 from 1, 3 and 6; entry 9 from 1 and 4; entry 10 from 1 and 5.

In the figure (Fig. 56), entry 1 is illustrated by conditions at B, D, F, and H; entry 2 at E and G, or any point between D and F or between F and H; entry 3 at A and C, or any point where $f'(x)$ lies below the axis OX; entry 4 at A and D, or between A and B, between C and E, between F and G; entry 5 between B and C, between E and F, between G and H, and to the right of H. Entry 6 is illustrated at B, C, E, F, and G; entry 7 at F; entry 8 at B; entry 10 at H. Entry 9 is not illustrated, but its significance is easily understood by analogy with 10.

Relations similar to those between these two curves exist between the graphs of $y' = f'(x)$ and $y'' = f''(x)$, or between the latter and $y''' = f'''(x)$. Use may thus be made of the successive derivatives of a function to determine its character. For example, corresponding to a point of inflexion of y, as at B, C, E, F, or G, there is a horizontal tangent to y', and therefore a zero point for y''. But y'' may be zero and y' have a horizontal tangent as at H, where there is no point of inflexion for y. At B, C, E, F and G the y'' curve would cross the axis OX, while at H it would be tangent to this axis.

Suppose it required to make a graph of the function

$$y = 0.29x^3 - 1.6x^2 + 1.56.$$

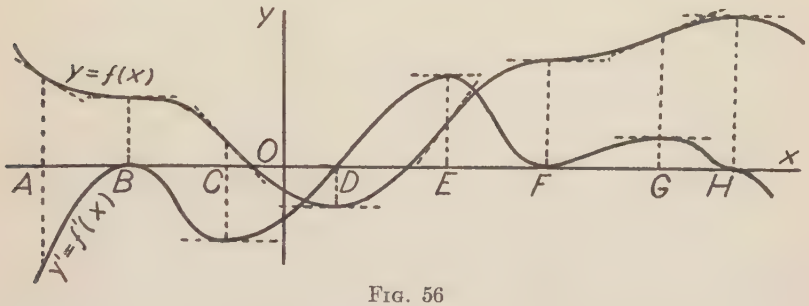


Fig. 56

Taking first and second derivatives we have

$$y' = 0.87x^2 - 3.2x$$

$$y'' = 1.74x - 3.2$$

Now the graph of y'' is a straight line (Fig. 57), with positive slope, cutting the x-axis at $x = 1.84$. At this value of x the graph of y' must have a horizontal tangent, and it must be concave upward at all points, by entries 1 and 4 in the table. Furthermore, $y' = -2.96$ when $x = 1.84$, and $y' = 0$ when $x = 0$ or $x = 3.68$, and is decreasing at 0, increasing at 3.68. We may make a sketch of y', when we see that y has horizontal tangents at $x = 0$ and $x = 3.68$, and that it is concave downward at the former and concave upward at the latter, while the point of inflexion is at $x = 1.84$. Finding the value of y at $x = 0$, $x = 1.84$ and $x = 3.68$, either by substitution or synthetic division, we are able to make a sketch of the graph, with exact location of these three important points.

2. **Roots of equations.** In solving an equation $f(x) = 0$ we may

gain knowledge of the positions of the roots by consideration of the equation $f'(x) = 0$. The following theorem is important.

Rolle's Theorem. If $f(x)$ is continuous on the interval from a to b, and if $f(a) = f(b) = 0$, and if $f(x) = 0$ has no real root between a and b, then if $f'(x)$ is continuous from a to b, the equation $f'(x) = 0$ has an odd number of real roots between a and b.

The truth of this is evident from consideration of the graph of $f(x)$. The function is zero at a and at b, and not between, so that if it increases (decreases) from a it must decrease (increase) toward b, so that the slope,

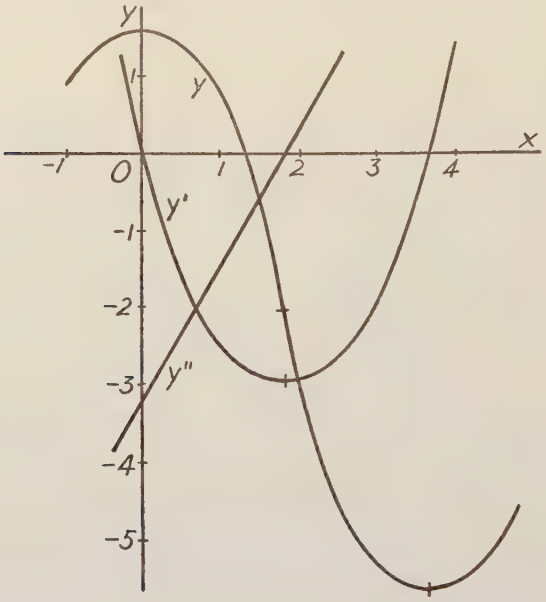


Fig. 57

$f'(x)$ must change sign an odd number of times, and being continuous, this means it must pass through zero. The theorem applies to transcendental equations as well as to algebraic equations. In the case of polynomial equations a convincing algebraic proof of the theorem is available.

Let $f(x) = (x - a)^r (x - b)^s \phi(x)$, where r and s are positive integers then

$$\frac{f'(x)}{f(x)} = \frac{r}{x - a} + \frac{s}{x - b} + \frac{\phi'(x)}{\phi(x)}.$$

or multiplying by $(x - a)(x - b)$,

$$\frac{(x-a)(x-b)f'(x)}{f(x)} = r(x-b) + s(x-a) + \frac{(x-a)(x-b)\phi'(x)}{\phi(x)}.$$

Now the right side is $r(a - b)$ when $x = a$ and $s(b - a)$ when $x = b$, and since these values are of opposite sign, the quantity must be zero at an odd number of places between a and b. Therefore $f'(x)$, the only factor of the left side which can be zero between a and b, must have an odd number of zero's on this interval.

If an equation $f(x) = 0$ contains one or more arbitrary constants or parameters the roots of the equation will, of course, depend on the values of these parameters. The parameters may take values which make two or more roots coincide, in which case they are called **double or multiple roots.** Since $f'(x) = 0$ has one real root between each two real roots of $f(x) = 0$, when two or more roots, say m roots, of the latter coincide at $x = a$, we must have $m - 1$ roots of the former coincide at $x = a$. That is, a root of multiplicity m of $f(x) = 0$ is a root of multiplicity $m - 1$ of $f'(x) = 0$. This, again, may be demonstrated algebraically for polynomial equations.

Let $f(x) = (x - a)^m \phi(x)$ then

$$f'(x) = (x - a)^{m-1}[m\phi(x) + (x - a)\phi'(x)]$$

from which the conclusion is obvious.

We note then that a double root of $f(x) = 0$ is a common root of this equation and $f'(x) = 0$, so that the elimination of x between these equations will give the relation among the coefficients or parameters necessary for coincident roots. The common factor of $f(x)$ and $f'(x)$ then equated to zero will determine the multiple root. The eliminant of $f(x) = 0$ and $f'(x) = 0$ is called the **discriminant** of $f(x) = 0$.

As an example let us find the discriminant of the cubic equation

$$x^3 + px^2 + qx + r = 0,$$

Taking the derivative with respect to x we have

$$3x^2 + 2px + q = 0,$$

and the Sylvester eliminant of these two equations is

$$\begin{vmatrix} 1 & p & q & r & 0 \\ 0 & 1 & p & q & r \\ 0 & 0 & 3 & 2p & q \\ 0 & 3 & 2p & q & 0 \\ 3 & 2p & q & 0 & 0 \end{vmatrix} = 0,$$

which gives

$$p^2q^2 - 4p^3r + 18\,pqr - 4q^3 - 27r^2 = 0.$$

If the coefficients fulfil this relation the cubic has a double root.

As another example, consider the family of lines $x \cos \alpha + y \sin \alpha = 5$, the angle α being the parameter of the family. These lines are all at distance 5 from the origin, and in general, there are two values of α that make the line go through a given point (x, y). For some points, however, the two values of α will coincide. To find these points we write the equation and its derivative with respect to α.

$$x \cos \alpha + y \sin \alpha = 5,$$

$$-x \sin \alpha + y \cos \alpha = 0,$$

and by squaring and adding we obtain

$$x^2 + y^2 = 25.$$

This is the equation of a circle, the locus of points for which the two values of α coincide, the curve to which all lines of the family are tangent.

EXERCISES XXXIV

1. Sketch the curve $12y = x^4 - 6x^2$, using derived functions to locate points of inflexion.

2. Use derivatives to aid in tracing the curve $y = x - x \log x$.

3. Given that the three roots of

$$x^3 - 6x^2 + 9x - 3 = 0$$

are real, show by the derivative that one is greater than 3, one between 1 and 3 and one less than 1.

4. By use of the derivative show that the discriminant of the quadratic equation $x^2 + px + q = 0$ is $p^2 - 4q$.

5. For what value of k does the equation $\cos \theta + \sin \theta = k$ have double solutions for θ?

6. The line $y = mx - 2pm - pm^3$ is normal to the parabola $y^2 = 4px$ for all values of m. In general, three normals pass through a point (x, y). Show that the locus of points for which two of these three normals coincide is the semi-cubical parabola $27py^2 = 4(x - 2p)^3$.

7. Show that the equation $f(x) = 0$ must have as many imaginary roots as has $f'(x) = 0$. Thus show that the equation

$$x^4 - 4x^3 + 15x^2 + 3x - 15 = 0$$

has two imaginary roots.

8. Show that the curve $y = 1 - x^{\frac{2}{3}}$ is concave upward for all values of x. Show that $y = 0$ when $x = 1$ and when $x = -1$, and is positive between these values. Why is the derivative not zero at some point between 1 and -1 according to Rolle's Theorem?

Art. 38. Maxima and Minima

If, at any value of x, say x_0, the graph of a function $y = f(x)$ has a horizontal tangent and is concave downward, then $y_0 = f(x_0)$ is larger than the value of y corresponding to any other x near x_0. This value y_0 is said to be a **maximum** value for y. Similarly, if the graph has, at x_0, a horizontal tangent and is concave upward, the corresponding value y_0 is lower than any neighboring value, and is called a **minimum** value of the function. By reference to the table in Art. 37, entries 1, 3 and 5, we see that we should have a maximum for y at any x for which $y' = 0$ and y'' is negative, and from entries 1, 2 and 4 that we should have a minimum for y at any x for which $y' = 0$ and y'' is positive. But we may have a maximum or minimum of y for values of x for which both y' and y'' are zero, as seen by entries 9 and 10 of the table.

By giving attention to the successive derivatives of a function $y = f(x)$ we may prescribe necessary and sufficient conditions for maxima and minima of the kind defined above. In case of either a maximum or minimum value of y we must have $y' = 0$. Suppose then that it is found that $y' = 0$ at x_0, x_1, x_2, etc. These values of x are called **critical values,** and we proceed to investigate each in turn, to see whether the corresponding value of y is a maximum, a minimum, or neither. Suppose that at x_0 we have $y = y_0, y' = 0, y'' = 0, \ldots y^{(n-1)} = 0, y^{(n)} = y_0^{(n)}$. That is, we compute values of the successive derivatives at $x = x_0$, until we come to one, the nth, that is not zero. This n may be 2 or any integer greater than 2. If $y_0^{(n)}$ is positive, $y^{(n-1)}$ is zero and increasing, $y^{(n-2)}$ is concave upward, then looking at entries 7 and 9 in the table of Art. 37, we see that $y^{(n-3)}$, $y^{(n-4)}$, etc., back to y', alternate between "point of inflexion" and "concave upward," all being tangent to the x-axis. Therefore if n is even, y_0 is a minimum, while if n is odd y has a "point of inflexion, rising." If $y_0^{(n)}$ is negative, we reason similarly from entries 8 and 10 of the table that y_0 is a maximum if n is

even and y has a "point of inflexion, falling," if n is odd. In more condensed form we have:

If at $x = x_0$ we have $y = y_0$, $y' = 0$, $y^{(n)} = y_0^{(n)}$, and all derivatives between y' and $y^{(n)}$ are zero, then if we have

(a) $y_0^{(n)} > 0$, n even, y_0 is a minimum for y;
(b) $y_0^{(n)} > 0$, n odd, y has rising point of inflexion;
(c) $y_0^{(n)} < 0$, n even, y_0 is a maximum for y;
(d) $y_0^{(n)} < 0$, n odd, y has falling point of inflexion.

For the function

$$y = 0.29x^3 - 1.6x^2 + 1.56$$

considered in Art. 37, the critical values are $x = 0$ and $x = 3.68$, and the former corresponds to a maximum of y and the latter to a minimum.

As a further illustration, consider the function

$$y = x^2 \sin x.$$

Here we have

$$y' = 2x \sin x + x^2 \cos x,$$

and this is zero when $x = 0$, or when $x = -2 \tan x$.
For the critical point $x = 0$, we have

$$y_0'' = [(2 - x^2) \sin x + 4x \cos x]_{x=0} = 0,$$

and

$$y_0''' = [-6x \sin x + (6 - x^2) \cos x]_{x=0} = 6.$$

In this case $n = 3$, odd, and $y^{(n)} = 6$, positive, so that y has a point of inflexion, rising, at $x = 0$.

Other critical values are solutions of the equation $x = -2 \tan x$, infinite in number. By the approximate methods of Art. 23, the least positive critical value is found to be at $x = 2.29$. At this value $\tan x = -1.14$, $\cos x = -.659$ and $\sin x = .752$. To find whether y'' is positive or negative at $x = 2.29$ we substitute $x = -2 \tan x$.

$y'' = (2 - 4 \tan^2 x) \sin x + 8 \tan x \cos x = \sin x(10 - 4 \tan^2 x)$.
Now since $\sin x$ is positive and $4 \tan^2 x$ is less than 10, y'' is positive at $x = 2\text{-}29$ and therefore

$$y_{x=2,29} = .752(2.29)^2 = 3.94$$

is a minimum value of y.

Attention will now be given to methods of treatment of problems in maxima and minima in the various forms in which they present themselves in practice.

1. **Explicit function of one variable.** If any quantity, say y, is a known function of an independent variable, say x, it may be desirable to know for what value of x the function y has its greatest or its least value. The range of values for x that require attention may be of indefinite extent, or it may lie within definite bounds. If the function is continuous, its greatest value may be a maximum of the kind discussed above, or it may be the value at one of the bounds of the range. Very rarely an extreme value of the function may occur at a value of x for which the derivative y′ is discontinuous, as at B or C in the figure (Fig. 58). At B the curve for y′ has a vertical asymptote, and is

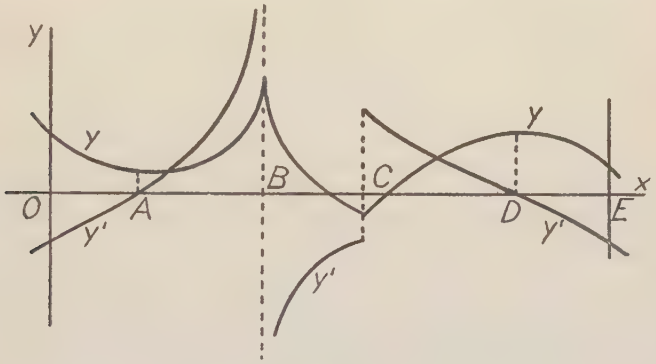


Fig. 58

positive to the left and negative to the right of it, showing that the y-curve has a sharp point, called a **cusp**. At C the curve for y′ has a finite discontinuity showing a sudden turn in the y-curve. To find the greatest value of y on the range from O to E we must compare its value at O, B and D. The maximum at D would be determined by the method set forth above. Similarly, to find the least value of y we would determine the minimum value at A, and compare this value with the values at C and at E.

The function

$$M = 3000x - 500x^2$$

considered in Art. 3 has a maximum at $x = 3$, where the derivative

$$\frac{dM}{dx} = 3000 - 1000x$$

is zero. At this place M = 4500. The minimum value of M is at the bound of the x-range, x = 8, where M = −8000. In this case, since a negative bending moment may produce as great a stress in the beam as a positive bending moment of the same magnitude, the two extreme values are to be compared numerically.

Again, suppose a certain quantity, y, to be given by the equation

$$y = \frac{2}{3}x - (x - 1)^{\frac{2}{3}}$$

where x has a range from 0 to 9. We have

$$y' = \frac{2}{3} - \frac{2}{3(x - 1)^{\frac{1}{3}}}.$$

This is zero when x = 2. The second derivative,

$$y'' = \frac{2}{9(x - 1)^{\frac{4}{3}}},$$

is positive at x = 2, hence the value of y at this place, $y = \frac{1}{3}$, is a minimum. The slope, y′ has an infinite discontinuity at x = 1, and changes from positive to negative, while y itself is finite, equal to 2/3, which gives a cusp. At x = 0, y = −1, and at x = 9, y = 2. The least value of the function on the range considered is therefore −1, at x = 0, and the greatest value is 2 at x = 9. A sketch of the graph of this function as shown here (Fig. 59) may be constructed from the data obtained.

2. **Implicit functions.** If y is a function of x defined by means of an equation of the form

$$f(x, y) = 0,$$

then the successive derivatives, y′, y″, etc., may be found by the method of Art. 14, and the critical points are found by solving the two equations,

$$f(x, y) = 0, \quad y' = 0,$$

for values x and y. These values may be substituted in the expressions for y″ and, if necessary, higher derivatives, and the criterion for maxima and minima is directly applicable.

If any values of x and y are found which satisfy the original equation f(x, y) = 0, and at which y′ is discontinuous, the value of y at this place

may be compared with the extreme values found otherwise to be the least and greatest values. A complete or adequate treatment of these unusual cases is beyond the scope of the present discussion, but is given further consideration in Chapter XII.

If the equation in the last illustration in 1 above, be rationalized we have

$$(2x - 3y)^3 - 27(x - 1)^2 = 0.$$

This is an implicit equation, and we derive

$$y' = \frac{2(2x - 3y)^2 - 18(x - 1)}{3(2x - 3y)^2}$$

$$y'' = \frac{2}{(2x - 3y)^2}$$

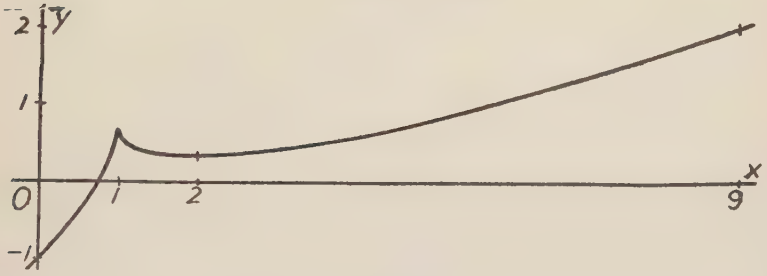


Fig. 59

Now y' will be zero if its numerator is zero and its denominator is not zero, so we write

$$(2x - 3y)^2 - 9(x - 1) = 0,$$

and solving this and the original equation for the two bracketed expressions we find, $x - 1 = 1$, $2x - 3y = 3$, from which $x = 2$ and $y = \frac{1}{3}$. For these values y'' is positive, so that $\frac{1}{3}$ is a minimum for y.

If $2x - 3y = 0$, then the original equation requires that $x - 1 = 0$, so that $x = 1$ and $y = \frac{2}{3}$. For these values the denominator of y', is zero, and while we cannot discuss the case completely, it is sufficient for present purposes to merely note these values of x and y. At the two bounds of the range for x, $x = 0$ and $x = 9$, the values of y are found by substitution in the original equation, and these values are

compared with the $\frac{1}{3}$ and $\frac{2}{3}$ to determine the least and greatest values.

3. **Auxiliary variables present.** Often, in expressing a quantity y as a function of x, the problem is simplified by the use of auxiliary variables. For example, suppose two such auxiliary variables, as u and v, are employed, then we should require three equations from which to eliminate u and v and express y as a function of x. Let the three equations be

$$f_1(x, y, u, v) = 0,$$

$$f_2(x, y, u, v) = 0,$$

$$f_3(x, y, u, v) = 0.$$

If, now we wish to find for what value of x there is a maximum or minimum for y, we may determine the critical values of x by forming three more equations by differentiating these three with respect to x, then we have six equations involving x, y, u, v, $\frac{dy}{dx}$, $\frac{du}{dx}$ and $\frac{dv}{dx}$. We may set $\frac{dy}{dx} = 0$, then eliminating u and v and their derivatives, solve for x and y. In cases of this kind, the critical values of x and the corresponding values of y, together with other known conditions of the problem, are usually sufficient, and to employ higher derivatives or to seek a discontinuity in $\frac{dy}{dx}$, would be neither convenient nor necessary.

Suppose it required to find the proper proportions for a cylindrical container with closed ends so that it shall have the least surface for a given volume. Let V denote volume, S surface, h altitude and r radius of the cylinder, then

$$V = \pi r^2 h, \qquad S = 2\pi rh + 2\pi r^2.$$

Here V is a known constant, and we may regard S as a function of r, with h as an auxiliary variable. Differentiating both equations with respect to r we have

$$0 = 2\pi rh + nr^2 \frac{dh}{dr},$$

$$\frac{dS}{dr} = 2\pi h + 2\pi r \frac{dh}{dr} + 4\pi 4.$$

Now putting $\frac{dS}{dr} = 0$, we have four equations from which to eliminate h

and $\frac{dh}{dr}$ and solve for r and S. In this case, however, the desired information is obtained by eliminating $\frac{dh}{dr}$ from the last two equations, obtaining $h = 2r$. We may, of course, use this relation with the first two equations to find the values of r, h and S in terms of the volume V.

Obviously, in the present case, it is not necessary to apply the remaining tests for minimum values, since the practical conditions of the problem assure us that we have found the correct answer to our question.

EXERCISES XXXV

1. Make a sketch of the curve

$$y = 2x^3 + 9x^2 - 60x + 12,$$

locating accurately the points of horizontal tangents, and points of inflexion.

2. Show that the catenary whose equation is $y = \frac{1}{2}(e^x + e^{-x})$ is tangent to the line $y = 1$ at $x = 0$, and that the curve lies entirely above this line.

3. Make a sketch of the curve

$$y = ax + b \cos x,$$

locating horizontal tangents and points of inflexion. (a) When $a = 1$, $b = 2$; (b) when $a = 2$, $b = 1$; (c) when $a = 1$, $b = 1$.

4. If a uniform beam of length L is supported at both ends and loaded uniformly, its deflection y at a distance x from one end is

$$y = K(x^4 - 2Lx^3 + L^3x),$$

where K is a constant. Find the greatest deflection, in terms of K, and show that the beam is concave upward at all points.

5. If instead of being merely supported the beam in 4 is firmly imbedded at both ends, the uniform load produces the deflection

$$y = K(x^4 - 2Lx^3 + L^2x^2).$$

Find, again, in terms of K, the greatest deflection, and find the position of points of inflexion.

6. A body of soldiers on an open plain is 8 miles west of a road which runs north and south. It is desired, in the least possible time, to reach

a point on the road 6 miles south of the present position of the soldiers. If they can march 4 miles per hour across the open field and 5 miles per hour on the road, toward what point of the road should they start?

7. Find the proportions of cylindrical container closed at both ends, so that it has the greatest capacity for a given surface.

8. The strength of a rectangular beam varies as the product of the breadth by the square of the depth. What is the form of the strongest beam that can be cut from a given circular log?

9. Show that the function ax^n has a point of inflexion at $x = 0$ if n is odd, and a maximum or minimum if n is even, the former if a is negative, the latter if a is positive.

10. The equation of the path of a projectile is

$$y = x \tan \theta - \frac{gx^2}{2v_o^2 \cos^2 \theta},$$

where θ is the angle of elevation, v_o is the initial velocity. What should be the value of θ in order that the projectile shall reach the greatest possible distance on the inclined plane given by $y = \frac{1}{4}x$, with a given v_o.

Art. 39. Osculating Circle, Curvature

1. **Osculating circle.** We have seen, in Art. 36, how to find the equation of a tangent line at any point of a curve whose equation is known. Now this is a line which goes through the required point and such that the value of $\frac{dy}{dx}$ computed from the equation of the line is the same as that computed from the equation of the given curve at the point in question. The two conditions, the required point and the required value of $\frac{dy}{dx}$, are sufficient to determine the line, since two independent conditions determine a line.

The general equation of a circle contains three arbitrary constants, and it requires three independent conditions to determine a circle. Having given the equation of a curve, let us write the equation of a circle which shall go through a given point of the curve, say (x_o, y_o), and such that the values of $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ at (x_o, y_o) computed from the equation of the circle shall have the same values as those computed from the equation of the given curve (Fig. 60).

Let us take for the equation of the circle

$$(x - a)^2 + (y - b)^2 = R^2,$$

where a, b and R are to be determined by the three conditions specified. If the circle goes through (x_o, y_o) we have

$$(x_o - a)^2 + (y_o - b)^2 = R^2$$

and if the first and second derivatives at (x_o, y_o), denoted by $\frac{dy_o}{dx_o}$ and $\frac{d^2y_o}{dx_o^2}$, are to have the required values we must have

$$2(x_o - a) + 2(y_o - b)\frac{dy_o}{dx_o} = 0,$$

$$2 + 2\left(\frac{dy_o}{dx_o}\right)^2 + 2(y_o - b)\frac{d^2y_o}{dx_o^2} = 0.$$

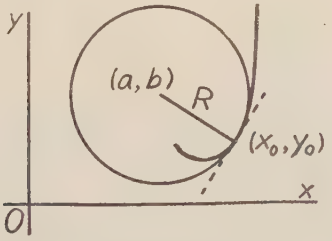


FIG. 60

Solving these three equations for a, b and R we obtain

$$a = x_o - \frac{\frac{dy_o}{dx_o}\left[1 + \left(\frac{dy_o}{dx_o}\right)^2\right]}{\frac{d^2y_o}{dx_o^2}}, \tag{1}$$

$$b = y_o + \frac{1 + \left(\frac{dy_o}{dx_o}\right)^2}{\frac{d^2y_o}{dx_o^2}}, \tag{2}$$

$$R = \frac{\left[1 + \left(\frac{dy_o}{dx_o}\right)^2\right]^{\frac{3}{2}}}{\frac{d^2y_o}{dx_o^2}}. \tag{3}$$

With these values of a, b and R the circle becomes the **osculating circle** at (x_o, y_o) to the curve from whose equation the values of the first and second derivatives are found.

2. **Curvature, rectangular coordinates.** From the straight line we obtain the notion of slope, or "rate of increase," since for the line this rate is constant. Similarly, from the circle we obtain the idea of **curvature,** or "rate of turning," since for the circle the rate of turning is constant. On the circle shown (Fig. 61), in going from A to B, a distance Δs along the arc, we turn from α to $\alpha + \Delta\alpha$, that is, we turn through the angle $\Delta\alpha$. The rate of turning is $\dfrac{\Delta\alpha}{\Delta s}$, and since $R\Delta\alpha = \Delta s$, we have, for the circle,

$$\text{Curvature} = \frac{\Delta\alpha}{\Delta s} = \frac{1}{R}.$$

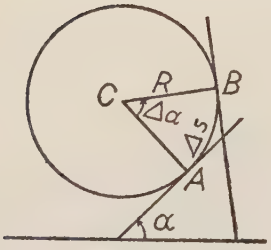


Fig. 61

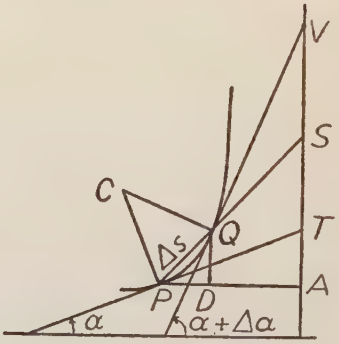


Fig. 62

Now, for any curve not a circle the curvature is not constant, but varies from point to point. Suppose a curve to be given by an equation in x and y, and let us consider curvature at a given point P of the curve (Fig. 62). Let the tangent lines at P and Q make respectively the angles α and $\alpha + \Delta\alpha$ with the x-axis. The angle turned through in going from P to Q is $\Delta\alpha$. In the next article (Art. 40), we shall give attention to the determination of length of curves in general, but for the present let us assume that any chord PQ subtends an arc of definite length Δs, and that the ratio of Δs to PQ approaches unity as Q approaches P, as in the case of a circle (see Art. 11). Corresponding to the increment Δs of the length we have the increment $PD = \Delta x$ in x and

$$\frac{\Delta x}{\Delta s} = \frac{PD}{PQ} \cdot \frac{PQ}{\Delta s} = \frac{PA}{PS} \cdot \frac{PQ}{\Delta s}$$

As Δs approaches zero the secant line PS approaches the tangent line PT, therefore

$$\frac{dx}{ds} = \frac{PA}{PT} = \cos \alpha. \tag{4}$$

The ratio of $\Delta\alpha$ to Δs, which is constant for a circle, depends on the value of Δs, and its limit as Δs approaches zero is the **curvature** of the curve at P. We have therefore,

$$\text{Curvature} = \frac{d\alpha}{ds} = \frac{d\alpha}{dx} \cdot \frac{dx}{ds} = \frac{d\alpha}{dx} \cos \alpha.$$

Now, $\tan \alpha = \dfrac{dy}{dx}$, so that

$$\frac{d\alpha}{dx} = \frac{d}{dx} \tan^{-1} \frac{dy}{dx} = \frac{\dfrac{d^2y}{dx^2}}{1 + \left(\dfrac{dy}{dx}\right)^2}$$

and

$$\cos \alpha = \frac{1}{\sqrt{1 + \tan^2 \alpha}} = \frac{1}{\sqrt{1 + \left(\dfrac{dy}{dx}\right)^2}}.$$

Therefore, for any curve

$$\text{Curvature} = \frac{d\alpha}{ds} = \frac{\dfrac{d^2y}{dx^2}}{\left[1 + \left(\dfrac{dy}{dx}\right)^2\right]^{\frac{3}{2}}}, \tag{5}$$

where the first and second derivatives are to be evaluated at the point for which the curvature is desired.

Comparison of this expression for curvature with that for the radius of the osculating circle shows, if we denote curvature by K, that

$$K = \frac{1}{R},$$

where R is the radius of the osculating circle. The curvature at any point of a curve, therefore, is that of the osculating circle at that point, and for this reason the osculating circle is called the **circle of curvature,**

its radius is the **radius of curvature,** and its center is the **center of curvature.**

3. **Curvature, polar coordinates.** In polar coordinates we have

$$\alpha = \theta + \psi,$$

where, as in Art. 36, θ is the vectorial angle and ψ is the angle between tangent line and radius vector. We have, therefore (Fig. 63),

$$K = \frac{d\alpha}{ds} = \frac{d\theta}{ds} + \frac{d\psi}{ds} = \left(\frac{d\theta}{d\rho} + \frac{d\psi}{d\rho}\right)\frac{d\rho}{ds}.$$

Now $\frac{d\rho}{ds}$, the limit of $\frac{\Delta\rho}{\Delta s}$, is $\cos\psi$, for referring to Figure 54,

$$\frac{\Delta\rho}{\Delta s} = \frac{QR}{\Delta s} = \frac{QR}{PR}\cdot\frac{PR}{\Delta s} = \frac{OR}{SR}\cdot\frac{PR}{\Delta s}.$$

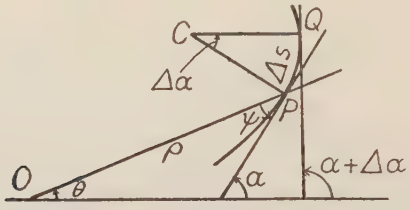


FIG. 63

Of these two factors the second approaches unity while the first approaches $\frac{OP}{TP}$, which is $\cos\psi$. Also since $\tan\psi = \rho\frac{d\theta}{d\rho}$, we have

$$\frac{d\psi}{d\rho} = \frac{\frac{d\theta}{d\rho} + \rho\frac{d^2\theta}{d\rho^2}}{1 + \rho^2\left(\frac{d\theta}{d\rho}\right)^2},$$

and

$$\cos\psi = \frac{1}{\sqrt{1+\tan^2\psi}} = \frac{1}{\sqrt{1+\rho^2\left(\frac{d\theta}{d\rho}\right)^2}},$$

therefore

$$K = \frac{2\frac{d\theta}{d\rho} + \rho^2\left(\frac{d\theta}{d\rho}\right)^3 + \rho\frac{d^2\theta}{d\rho^2}}{\left[1+\rho^2\left(\frac{d\theta}{d\rho}\right)^2\right]^{\frac{3}{2}}}. \qquad (6)$$

If the form of the polar equation of a curve makes it desirable to regard θ as the independent variable, then we may use a form for K obtained by using the identical relations

$$\frac{d\theta}{d\rho} = \frac{1}{\frac{d\rho}{d\theta}}, \qquad \frac{d^2\theta}{d\rho^2} = \frac{-\frac{d^2\rho}{d\theta^2}}{\left(\frac{d\rho}{d\theta}\right)^3},$$

by which we have after substituting and reducing,

$$K = \frac{\rho^2 + 2\left(\frac{d\rho}{d\theta}\right)^2 - \rho\frac{d^2\rho}{d\theta^2}}{\left[\rho^2 + \left(\frac{d\rho}{d\theta}\right)^2\right]^{\frac{3}{2}}}. \tag{7}$$

EXERCISES XXXVI

1. Show that the osculating circle at any point (x, y) of the parabola $y^2 = 4px$ has a radius equal to

$$\frac{(y^2 + 4p^2)^{\frac{3}{2}}}{4p^2}.$$

Find the center of the osculating circle for the point on the parabola where $x = p$. Also for the vertex. (Note that the first and second derivatives, y' and y'', are infinite at the vertex, and the expressions for a, b, and R must be simplified before evaluating for this particular point.)

2. Show that, for any point of the hyperbola $xy = k$, the radius of curvature is $\frac{d^3}{2k}$, where d is the distance from the point to the origin.

3. Find the curvature of the logarithmic spiral $\rho = e^{a\theta}$ at the point where $\theta = 0$. Show that for any point of this curve the radius of curvature is $R = \rho\sqrt{1 + a^2}$.

4. For the cycloid

$$x = a(\theta - \sin\theta),\ y = a(1 - \cos\theta),$$

show that the radius of curvature at any point is twice the length of the normal, and show that this is zero at $\theta = 0$, and at $\theta = 2\pi$.

5. Using the identical relations $x = \rho\cos\theta$, $y = \rho\sin\theta$, derive the expression for K in polar coordinates by substituting in the expression in rectangular coordinates the values of both $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ in terms of ρ, θ, $\frac{d\rho}{d\theta}$ and $\frac{d^2\rho}{d\theta^2}$.

6. Find the radius of curvature at any point (x, y) of the sine curve $y = \sin x$, and show that the curvature is zero when $y = 0$.

7. Find an expression for curvature in terms of t for the curve

$$x = a \cos^3 t, \qquad y = a \sin^3 t.$$

8. Show that the area of the osculating circle at any point of the cardioid $\rho = \frac{9}{8\pi}(1 - \cos \theta)$ is equal to the radius vector of the point.

9. Show that the curvature at a point of inflexion on any curve is zero.

10. Find the center of curvature of the parabola $y^2 = 4px$ at a general point (x_o, y_o). Show by eliminating x_o and y_o that the locus of the center of curvature is the semi-cubical parabola to which all normals to the given parabola are tangent. (See exercise 6 of XXXIV.)

CHAPTER IX

APPLICATIONS OF INTEGRATION

Art. 40. Length of Arc of a Curve

It is necessary to see just what is meant by "Length of arc of a curve." The length of a segment of a straight line, or the distance between two points, is a definite concept, and in terms of this we must define the idea of length of arc.

Consider a curve (Fig. 64) joining two points, A and B, and let a set of points on the curve between A and B be so chosen that the chords joining successive points are each of length less than or equal to h.

Fig. 64

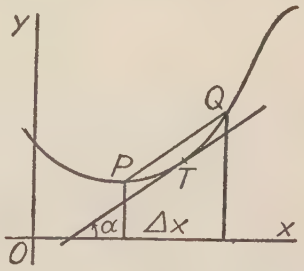


Fig. 65

The sum of the lengths of these chords is a definite quantity, say L. As the number of points between A and B is indefinitely increased, the value of h approaching zero as a limit, the value L may approach a definite limit, s, and this limit may be independent of the particular points chosen between A and B. If L has this limit s, however the points are chosen, providing only that h approaches zero, then **this limit of L is the length of the arc from A to B**; otherwise, the notion of length is not definite for the given curve.

The conditions under which the quantity L will have a definite limit cannot be established, or even stated precisely here. It is sufficient for our present purpose to know that the definition as given yields a definite value for the length of arc of a curve joining two points if the curve is continuous, does not intersect itself, and has a definite tangent line at all points. If the curve intersects itself or fails to have a definite

tangent line at a definite number of points, lengths of portions between such exceptional points are definite, and may be added to obtain the length of the whole.

Suppose the equation of a continuous curve with definite tangent line at each point to be given as $f(x, y) = 0$. Let s denote the length of the arc of the curve from a certain point (x_o, y_o) to any point (x, y). We wish to derive an expression for s in terms of the variables x and y. Let P and Q (Fig. 65) be points on the curve. The slope of the tangent line varies continuously as the point of tangency moves from P to Q, and at some points is greater than the slope of PQ and at some points less, therefore, by the general theorem at the end of Art 7, there is some point, say T, between P and Q at which the tangent line has a slope equal to that of PQ. If α is the angle made by the tangent at T with the x-axis, and Δx is the difference of abscissas of P and Q, then

$$PQ = \sec \alpha \, \Delta x = \sqrt{1 + \tan^2 \alpha} \, \Delta x = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \Delta x.$$

Now the limit of the sum of such elements from x_o to x is the length required, therefore

$$s = \int_{x_o}^{x} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \, dx, \tag{1}$$

where $\frac{dy}{dx}$ is to be expresed as a function of x by means of the equation of the curve.

By the symmetry of the situation with respect to x and y, an equivalent expression is

$$s = \int_{y_o}^{y} \sqrt{1 + \left(\frac{dx}{dy}\right)^2} \, dy, \tag{2}$$

where $\frac{dx}{dy}$ is to be expressed as a function of y by means of the equation of the curve.

If the curve is given by parametric equations, as

$$x = f_1(t), \qquad y = f_2(t),$$

then we obtain very obviously, from either (1) or (2),

$$s = \int_{t_o}^{t} \sqrt{[f'_1(t)]^2 + [f'_2(t)]^2} \, dt, \tag{3}$$

where $x = x_o$, $y = y_o$ when $t = t_o$. It may be noted here that if x and y give the position of a particle in the plane, and if t denotes time, then $f'_1(t)$ and $f'_2(t)$ are the components of velocity in the directions of the axes of x and y, and the integrand is the speed, v, of the particle along the path. The integral is, then

$$s = \int_{t_o}^{t} v \, dt, \tag{3'}$$

the distance traversed along the path from the time t_o to the time t.

If the equation is given in polar coordinates (ρ, θ), more convenient forms are derived from (1) and (2) by means of the identical relations

$$x = \rho \cos \theta, \qquad y = \rho \sin \theta.$$

The forms are:

$$s = \int_{r_o}^{r} \sqrt{1 + \rho^2 \left(\frac{d\phi}{d\rho}\right)^2} \, d\rho, \tag{4}$$

$$s = \int_{\theta_o}^{\theta} \sqrt{\left(\frac{d\rho}{d\theta}\right)^2 + \rho^2} \, d\theta, \tag{5}$$

the integrand in 4 to be expressed as a function of ρ, and that in (5) as a function of θ by means of the equation of the curve.

EXERCISES XXXVII

1. Find the length of the parabola $y^2 = 16x$ from the vertex to the point (4, 8).

2. Find the length of the catenary $y = \frac{a}{2}\left(e^{\frac{x}{a}} + e^{-\frac{x}{a}}\right)$ from $x = 0$ to the point (x, y).

Ans. $\frac{a}{2}\left(e^{\frac{x}{a}} - e^{-\frac{x}{a}}\right)$.

3. Find the length of one arch of the cycloid $x = a(\theta - \sin \theta)$, $y = a(1 - \cos \theta)$.

Ans. 8a.

4. If a particle moves so that its position at time t is given by

$$x = \frac{1}{2}t^2, \qquad y = \frac{2}{3}t^3,$$

show that it moves a distance of $\frac{17^{\frac{3}{2}} - 1}{12}$ units along the path in two seconds from rest at $t = 0$.

5. Find the perimeter of the cardioid $\rho = a(1 + \cos\theta)$.

Ans. 8a.

6. Find the length of the spiral $\rho = e^{a\theta}$ from $\rho = 0$ to $\rho = a$.

Ans. $\sqrt{a^2 + 1}$.

7. Find by integration in both polar and rectangular coordinates the circumference of a circle of radius R.

Art. 41. Areas and Volumes

1. **Plane areas, rectangular coordinates.** We have become familiar with the idea that an integral of the form

$$\int_{x_0}^{x} f(x)\,dx,$$

represents the area between the curve $y = f(x)$ and the x-axis from x_0 to x, and that the differential of this area is

$$dA = ydx = f(x)dx.$$

Now it will be helpful to think of this area as being generated by the motion of the vertical line of length y, fixed in direction and varying in length. The differential of area is then thought of as the length $f(x)$ of the line multiplied by the differential, dx, of the distance through which it moves. This is analogous to the thought of a curve as generated by a moving point, the differential of length being the differential, ds, of the distance through which the point moves.

From this point of view it is readily recognized that the differential of any plane area may be denoted by the expression,

$$dA = \phi(x)\,dx,$$

where $\phi(x)$ represents as a function of x the distance across the area at right angles to the x-axis, as PQ in the figure (Fig. 66). The area is the value of the integral

$$A = \int_{x_1}^{x_2} \phi(x)\,dx, \tag{1}$$

where x_1 and x_2 are the extreme values of x on the area.

Similarly the area is given by

$$A = \int_{y_1}^{y_2} \theta(y)\, dy, \tag{2}$$

where $\theta(y)$ represents as a function of y the distance across the area at right angles to the y-axis, and y_1 and y_2 are the extreme values of y on the area.

For example, let us find the area bounded by the parabola $y^2 = 4x$ and the line $y = -x + 3$. Sketching the curve and the line (Fig. 67), and solving the two equations for the x and y of the points of intersection, we find that from $x = 0$ to $x = 1$ the distance across the area

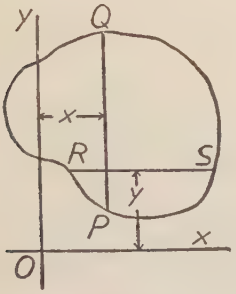


Fig. 66

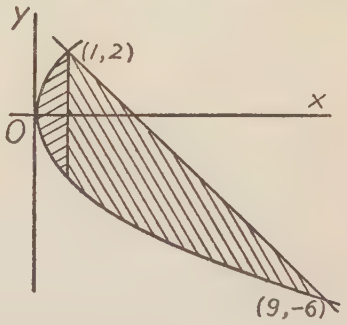


Fig. 67

vertically is $4\sqrt{x}$, while from $x = 1$ to $x = 9$ the distance across is $-x + 3 + 2\sqrt{x}$, so that the total area is

$$A = 4\int_0^1 \sqrt{x}\, dx + \int_1^9 (-x + 3 + 2\sqrt{x})\, dx = \frac{64}{3}.$$

But the distance across the area horizontally is $3 - y - \frac{1}{4}y^2$, and therefore

$$A = \int_{-6}^{2} (3 - y - \frac{1}{4}y^2)\, dy = \frac{64}{3}.$$

2. Plane areas, polar coordinates. It is frequently desirable to regard an area as generated by the motion of a varying portion of the arc of an expanding circle (Fig. 68). If the length of arc intercepted by the boundaries of the area is expressed as a function $f(\rho)$ of the radius ρ of

the expanding circle, then the differential of area may be expressed as

$$dA = f(\rho)\, d\rho,$$

and the area is given by the integral

$$A = \int_{\rho_1}^{\rho_2} f(\rho)\, d\rho, \tag{3}$$

where ρ_1 and ρ_2 are the extreme values of ρ on the area.

Again, an area may be regarded as generated by a variable portion of a radius vector as it swings about a fixed point, O. Let P and Q be the points (Fig. 69) at which the radius vector cuts the boundaries of the

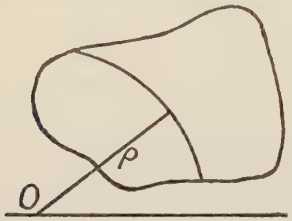


Fig. 68

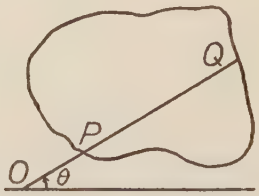


Fig. 69

curve, then the differential of area is the length PQ times the differential of distance the middle point of PQ moves, which is

$$\frac{1}{2}(OP + OQ)\, d\theta,$$

That is

$$dA = \frac{1}{2}(OQ - OP)(OP + OQ)\, d\theta = \frac{1}{2}\left(\overline{OQ}^2 - \overline{OP}^2\right)d\theta.$$

If the quantity $\frac{1}{2}\left(\overline{OQ}^2 - \overline{OP}^2\right)$ is expressed as a function, $f(\theta)$ of the vectorial angle θ, then the area is given by the integral

$$A = \int_{\theta_1}^{\theta_2} f(\theta) d\theta. \tag{4}$$

To illustrate, let us find the area cut from a circle of radius a by a line at distance b from its center (Fig. 70). The length of arc intercepted by the boundaries at distance ρ, from the center is $2\rho\sec^{-1}\frac{\rho}{b}$, and the extreme values of ρ are b and a.

The area is

$$A = \int_b^a 2\rho \sec^{-1}\frac{\rho}{b} d\rho = a^2 \sec^{-1}\frac{a}{b} - b\sqrt{a^2 - b^2}.$$

Again, the distances OQ and OP are respectively a and b sec θ, so that, integrating in terms of θ.

$$A = \int_{-\sec^{-1}\frac{a}{b}}^{\sec^{-1}\frac{a}{b}} \frac{1}{2}(a^2 - b^2 \sec^2\theta)\, d\theta = a^2 \sec^{-1}\frac{a}{b} - b\sqrt{a^2 - b^2}$$

3. **Area of surfaces of revolution.** A surface of revolution is generated by revolving a curve, as y = f(x), about an axis, as Ox (Fig. 71). We may also regard this surface as generated by the motion of a circle

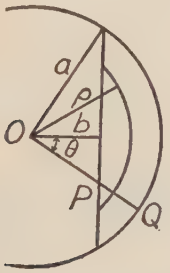


FIG. 70

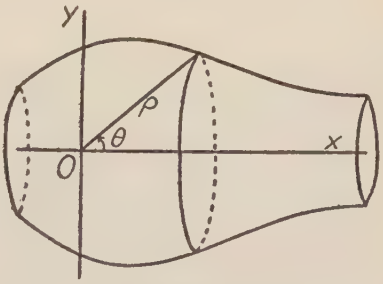


FIG. 71

of varying radius, the plane of the circle remaining perpendicular to Ox, and its radius always equal to the variable y. The differential of area is, then, the circumference of the moving circle times the differential of distance ds that the circle moves along the curve, and the area is given by

$$A = \int_{s_1}^{s_2} 2\pi y\, ds. \tag{5}$$

To perform the integration, the integral is expressed in one of the forms

$$A = \int_{x_1}^{x_2} 2\pi f(x) \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx, \tag{6}$$

$$A = \int_{y_1}^{y_2} 2\pi y \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy. \tag{7}$$

In the former of these the integrand is expressed as a function of x, in the latter as a function of y.

If the curve were given in polar coordinates, the initial line being the axis of revolution, then the differential area takes the form

$$dA = 2\pi\rho \sin\theta \, ds,$$

and if $\rho = f(\theta)$ is the equation of the curve we have

$$A = \int_{\theta_1}^{\theta_2} 2\pi \, f(\theta) \sin\theta \sqrt{\rho^2 + \left(\frac{d\rho}{d\theta}\right)^2} \, d\theta, \tag{8}$$

while if the curve is given in the form $\theta = f(\rho)$, a convenient form for the integral would be

$$A = \int_{\rho_1}^{\rho_2} 2\pi\rho \sin[f(\rho)] \sqrt{1 + \rho^2\left(\frac{d\theta}{d\rho}\right)^2} \, d\rho. \tag{9}$$

To find the area of a sphere of radius a, we may consider the surface formed by revolving the circle $x^2 + y^2 = a^2$ about the x-axis. We have the alternative forms

$$A = \int_{-a}^{a} 2\pi\sqrt{a^2 - x^2} \cdot \sqrt{1 + \frac{x^2}{a^2 - x^2}} \, dx = 2\pi \int_{-a}^{a} a \, dx = 4\pi a^2;$$

$$A = 2\int_{0}^{a} 2\pi y \sqrt{1 + \frac{y^2}{a^2 - y^2}} \, dy = 4\pi a \int_{0}^{a} \frac{dy}{\sqrt{a^2 - y^2}} = 4\pi a^2.$$

To treat this problem by polar coordinates it will be convenient to consider the circle $\rho = 2a \cos\theta$, rotated about the initial line. Then we have the alternative forms

$$A = \int_{0}^{\frac{\pi}{2}} 2\pi \cdot 2a \cos\theta \sin\theta \sqrt{4a^2\cos^2\theta + 4a^2\sin^2\theta} \, d\theta$$

$$= 8\pi a^2 \int_{0}^{\frac{\pi}{2}} \cos\theta \sin\theta \, d\theta = 4\pi a^2;$$

$$A = \int_{0}^{2a} 2\pi\rho \sqrt{1 - \frac{\rho^2}{4a^2}} \sqrt{1 + \frac{\rho^2}{4a^2 - \rho^2}} \, d\rho = 2\pi \int_{0}^{2a} \rho \, d\rho = 4\pi a^2.$$

These alternative forms are written here to illustrate their use. In actual work we use that one of the four which promises to give results most readily.

4. **Volumes.** We may regard a volume as being generated by the motion of a varying portion of a plane, the plane remaining perpendicular to a given fixed line. The differential of volume is the area of the moving plane times the differential of the distance moved along the fixed line. If the fixed line is the x-axis, and the area of the moving plane, the area of a cross-section of the volume (Fig. 72), is expressed as a function, f(x), of x, the volume is given by

$$V = \int_{x_1}^{x_2} f(x)\, dx. \tag{10}$$

As a special case, if the volume is bounded by a surface of revolution formed by revolving a curve $y = \phi(x)$ about the x-axis, then the area of the moving plane is πy^2, and the integral becomes

$$V = \pi \int_{x_1}^{x_2} [\phi(x)]^2\, dx. \tag{11}$$

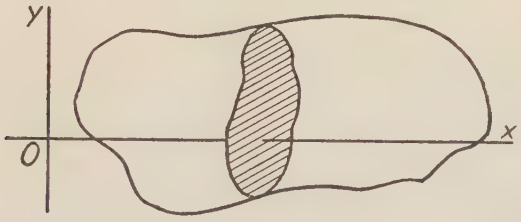


FIG. 72

Very frequently it is convenient to regard a volume as generated by a varying portion of the surface of an expanding sphere. If the portion of the surface of the sphere of radius ρ that lies in the given volume is expressed as a function of ρ, say $f(\rho)$, then the volume may be expressed as

$$V = \int_{\rho_1}^{\rho_2} f(\rho)\, d\rho. \tag{12}$$

For the volume of a sphere of radius a, using (11) we have

$$V = \pi \int_{-a}^{a} (a^2 - x^2) dx = \frac{4}{3}\pi a^3,$$

or, using (12) we may write, taking pole at center,

$$V = \int_{0}^{a} 4\pi\rho^2 d\rho = \frac{4}{3}\pi a^3.$$

EXERCISES XXXVIII

1. Show by integration that the area of the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ is πab.

2. Find the area between the parabola $x^2 = 4ay$ and the witch $y = \dfrac{8a^3}{x^2 + 4a^2}$. Ans. $\left(2\pi - \dfrac{4}{3}\right)a^2$.

3. Find the area of one arch of the cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$. Ans. $3\pi a^2$.

(Suggestion: Use form (1), but express $\phi(x)$ and dx in terms of θ and $d\theta$, and use limits of θ.)

4. Show that the entire area of the cardioid $\rho = a(1 - \cos\theta)$ is $\dfrac{3}{2}\pi a^2$.

(Use form (4), observing that OP is zero.)

5. Use form (8) to determine the area of the surface generated by revolving the cardiod of example 4 about the initial line.

$$\text{Ans. } \frac{32}{5}\pi a^2,$$

6. Find the volume bounded by the surface in example 5.

$$\text{Ans. } \frac{8}{3}\pi a^3.$$

(Use form (11), but express y, or $\phi(x)$, and dx in terms of θ $d\theta$ by means of the relations $x = \rho\cos\theta$, $y = \rho\sin\theta$, and the equation of the cardiod.)

7. Apply form (12) to the problem stated in example 6.

8. If a beam of rectangular section is fixed at one end and loaded at the other, the greatest tensile stress in the section at distance x from the load is given by

$$f = \frac{6Wx}{bh^2}$$

where f is stress in pounds per square inch, W is weight in pounds, x is distance in inches, and b and h are width and depth of the section in inches. If b and h are so related that the stress f is the same for all values of x the beam is said to be of "uniform strength." If $f = 10{,}000$ for all values of x, and $W = 20{,}000$, and the total length of the beam is 36 inches, find the volume of the beam in cubic inches under each of the following conditions.

(a) The width b is uniformly 3 inches.

(b) The depth h is uniformly 3 inches.

(c) The section is of uniform shape, $b = \frac{3h}{16}$.

9. In a cylindrical tank with horizontal axis, 15 feet radius and 60 feet long, oil stands 6 feet deep. How many cubic feet of oil are required to bring the depth to 13 feet?

10. A solid is described by revolving the curve $y = a \log_e bx$ about the y-axis. Find the volume below the x-axis. If this solid were suspended, and if the weight per unit volume is w, show that the weight below any horizontal cross-section is proportional to the area of the section, and that the average stress per unit area due to the weight is $\frac{1}{2}$ aw.

Art. 42. Moments and Mean Values

We have seen how to express differentials of arc, of area and of volume, and how to obtain totals of these quantities by integration. If, before integrating, the differential element is multiplied by a power of its distance from a fixed point, line or plane the result of integrating is called a moment of arc, area or volume with respect to this point, line or plane.

More definitely, if the differential of arc, area or volume is expressed as $f(x)\,dx$, then the integral

$$\int_{x_0}^{x_1} (x - a)^n f(x)\,dx) \tag{1}$$

gives the **nth moment with respect to the locus x = a** of that portion of the arc, area or volume between x_0 and x_1. A similar statement may be made putting y for x, and if the differential is expressed as $f(\rho)\,d\rho$, then the integral

$$\int_{\rho_1}^{\rho_2} \rho^n f(\rho)\,d\rho \tag{2}$$

gives the **nth polar moment with respect to the pole** of that part of the arc, area or volume between ρ_1 and ρ_2.

In applying (1) we are thinking of an arc generated by a moving point, of an area generated by a moving vertical line, of a surface of revolution generated by a circle whose plane is perpendicular to the x-axis, or of a volume generated by a plane area perpendicular to the x-axis. In any of these cases x-a denotes the distance from the locus

$x = a$, which is uniform for the whole moving element. Similarly, in applying (2) we are thinking of an arc as generated by a moving point, of an area by a moving circular arc of a surface of revolution generated by a circle whose plane is perpendicular to the initial line, or of a volume generated by a portion of spherical surface. In any case ρ denotes the distance of the whole moving element from the pole. Variations of this method of procedure will appear in the illustrations below, and a further discussion of the subject is given in Art. 44.

The mean value of x^n on the range (arc, area or volume) given by the integral

$$\int_{x_0}^{x_1} f(x)\, dx$$

may be denoted by $(x^n)_m$, and is defined by the equation

$$(x^n)_m \int_{x_0}^{x_1} f(x)\, dx = \int_{x_0}^{x_1} x^n f(x)\, dx. \tag{3}$$

Similarly, the mean value of ρ^n for the range indicated is defined by the equation

$$(\rho^n)_m \int_{\rho_0}^{\rho_1} f(\rho)\, d\rho = \int_{\rho_0}^{\rho_1} \rho^n f(\rho)\, d\rho. \tag{4}$$

In practical applications the most important moments are given in the cases $n = 1$ and $n = 2$. The first moment, $n = 1$, is sometimes called the **statical moment,** because of its relation to equilibrium of systems of parallel forces, and the mean value of x for a given range is the abscissa of the **center of gravity** of the range. Other coordinates of the center of gravity are similarly defined. The second moment, $n = 2$, is called the **moment of inertia,** because of its relation to inertia forces in rotating bodies, and the mean value of x^2, y^2 or ρ^2 is called the squared **radius of gyration** with respect to the point, line or plane from which distance is measured.

The significance of these definitions and the methods of applying them will be made clear by a few examples.

Consider the arc of the parabola $y = x^2$ between the points (1, 1) and (3, 9). The length of this arc is given with equal exactness by the two integrals

$$s = \int_1^3 \sqrt{1 + 4x^2}\, dx = \int_1 \sqrt{1 + \frac{1}{4y}}\, dy.$$

The first moment of the arc with respect to the y-axis, which we may denote by M_y, is given by

$$M_y = sx_m = \int_1^3 x\sqrt{1 + 4x^2}\,dx = \int_1^9 y^{\frac{1}{2}}\sqrt{1 + \frac{1}{4y}}\,dy.$$

The last integral written here is an alternative form which is available in the case of the arc because for all points on the arc the relation $y = x^2$ holds. Similarly for the first moment with respect to the x-axis we have the two forms

$$M_x = sy_m = \int_1^9 y\sqrt{1 + \frac{1}{4y}}\,dy = \int_1^3 x^2\sqrt{1 + 4x^2}\,dx.$$

The center of gravity of the arc is the point (x_m, y_m) found by these equations. The second moments, or moments of inertia, with respect to the y and x axis may be denoted by I_y and I_x respectively, and are expressed by the equations

$$I_y = s(x^2)_m = \int_1^3 x^2\sqrt{1 + 4x^2}\,dx = \int_1^9 y\sqrt{1 + \frac{1}{4y}}\,dy,$$

$$I_x = s(y^2)_m = \int_1^9 y^2\sqrt{1 + \frac{1}{4y}}\,dy = \int_1^3 x^4\sqrt{1 + 4x^2}\,dx.$$

It may be noted that the condition $y = x^2$, in this example gives us the special relation $M_x = I_y$, a relation not generally true.

Now let us consider a rectangle of base a and altitude b, and let us choose the base as the x-axis and the left hand edge as the y-axis. The area is

$$A = \int_0^a b\,dx = \int_0^b a\,dy = ab.$$

The first moments with respect to the edges are

$$M_y = Ax_m = \int_0^a b\,x\,dx = \frac{1}{2}a^2b,$$

$$M_x = Ay_m = \int_0^b ay\,dy = \frac{1}{2}ab^2.$$

and from this we have $x_m = \frac{a}{2}$, $y_m = \frac{b}{2}$ so that the center of gravity is at

the center of the rectangle. The moments of inertia with respect to the edges are

$$I_y = A(x^2)_m = \int_0^a b\, x^2\, dx = \frac{1}{3}\, a^3 b,$$

$$I_x = A\,(y^2)_m = \int_0^b a\, y^2\, dy = \frac{1}{3}\, ab^3,$$

from which we have $(x^2)_m = \frac{a^2}{3}$, $(y^2)_m = \frac{b^2}{3}$.

As a third example, consider the area under the curve $y = \sin x$ from $x = 0$ to $x = \pi$. The area is

$$A = \int_0^\pi \sin x\, dx = 2,$$

and the first moment with respect to the y-axis is

$$M_y = Ax_m = \int_0^\pi x \sin x\, dx = \pi$$

showing $x_m = \frac{\pi}{2}$. The first moment with respect to the x-axis may be written.

$$M_x = Ay_m = \int_0^1 y(\pi - 2 \sin^{-1} y)\, dy,$$

where the value of $\sin^{-1} y$ is given its first quadrant value, but it is more convenient to take the element of area as $\sin x\, dx$, and multiply this rectangle by the mean value of y on it; namely, $\frac{\sin x}{2}$ as shown by previous example. This gives

$$M_x = Ay_m = \int_0^\pi \frac{\sin^2 x}{2}\, dx = \frac{\pi}{4}.$$

so that $y_m = \frac{\pi}{8}$. For the moments of inertia we have

$$I_y = A(x^2)_m = \int_0^\pi x^2 \sin x\, dx = \pi^2.$$

$$I_x = A(y^2)_m = \int_0^\pi \frac{\sin^3 x}{3}\, dx = \frac{4}{9}.$$

the last integral being the result of multiplying the elementary rectangle sin x dx by the mean value of y^2 on it, $\frac{\sin^2 x}{3}$, as given by the previous example.

Next, consider a sector of a circle of radius r, and let the angle of the sector be α. Take the pole at the center, then the area of the sector is

$$A = \int_0^r \alpha\rho d\rho = \frac{\alpha r^2}{2},$$

and the first polar moment is

$$M_o = A\rho_m = \int_0^r \alpha\rho^2 d\rho = \frac{\alpha r^3}{3}$$

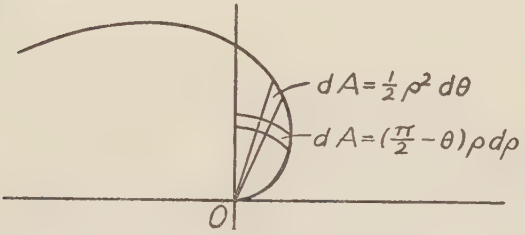


FIG. 73

showing that the mean value of ρ is $\rho_m = \frac{2}{3}r$. The polar moment of inertia of the sector is

$$I_o = A(\rho^2)_m = \int_0^r \alpha\rho^3 d\rho = \frac{\alpha r^4}{4}$$

showing that the mean value of ρ^2 is $(\rho^2)_m = \frac{r^2}{2}$.

As a final example, consider the area swept over by the radius vector to the curve $\rho = a\theta$ from $\theta = 0$ to $\theta = \frac{\pi}{2}$. For area we have two forms (Fig. 73).

$$A = \int_0^{\frac{a\pi}{2}} \left(\frac{\pi}{2} - \frac{\rho}{a}\right)\rho d\rho = \int_0^{\frac{\pi}{2}} \frac{a^2\theta^2}{2} d\theta = \frac{a^2\pi^3}{48}.$$

For the first polar moment we have, using the first form for dA,

$$M_o = A\rho_m = \int_0^{\frac{a\pi}{2}} \left(\frac{\pi}{2} - \frac{\rho}{a}\right)\rho^2 d\rho = \frac{a^3\pi^4}{192}.$$

showing that $\rho_m = \frac{a\pi}{4}$. But in this case a more convenient form may be obtained by taking the element of area in the second form, and, being in the form of a sector, the mean value of ρ on this element is $\frac{2}{3}a\theta$ according to the previous example, so that the integral becomes

$$M_o = A\rho_m = \int_o^{\frac{\pi}{2}} \frac{a^3\theta^3}{3}\, d\theta = \frac{a^3\pi^4}{192}.$$

The polar moment of inertia is most easily found, in this case, by taking the element of area in the sector form, and multiplying by the mean value of ρ^2 on this sector, $\frac{a^2\theta^2}{2}$, obtaining

$$I_o = A(\rho^2)_m = \int_o^{\frac{\pi}{2}} \frac{a^4\theta^4}{4}\, d\theta = \frac{a^4\pi^5}{640},$$

from which we have

$$(\rho^2)_m = \frac{3}{40}a^2\pi^2.$$

EXERCISES XXXIX

1. Find the center of gravity of the arc of the parabola $y^2 = 8x$ from the origin to the point (2, 4).

2. Find the center of gravity of the area of a triangle whose sides are 3, 4, 5.

3. Find the moment of inertia with respect to the x-axis of the area cut from the first quadrant by the line $2x + 3y = 30$.

4. Find the center of gravity of a semi-circular arc. Find the moment of inertia of a complete circular arc with respect to a diameter.

5. Find the first and second polar moments and the radius of gyration of the area bounded by the curve $\rho = e^\theta$, the initial line and the radius vector to the point $\left(e^{\frac{\pi}{4}}, \frac{\pi}{4}\right)$.

6. Show that the polar moment of inertia of a circular area with respect to the center is twice the moment of inertia with respect to a diameter.

7. A pyramid has a rectangular base, dimensions a and b, and an

altitude h. Find the center of gravity of the volume and the moment of inertia with respect to the base plane.

8. Find the polar moment of inertia of a solid sphere with respect to its center.

9. Find the center of gravity of the area under one arch of the cycloid $x = a(\phi - \sin \phi)$, $y = a(1 - \cos \phi)$.

10. Given a right circular cone, radius of base r, altitude h, find the center of gravity of the curved surface, also of the volume. Find the moment of inertia of the curved surface and of the volume, both with respect to the axis and with respect to the plane of the base.

Art. 43. Distributed Quantities: Masses, Forces

If a quantity is distributed over a range (arc, area, volume) the intensity of the quantity may vary from point to point, and the value of this intensity at a point may be expressed as a function of the coordinates of the point. This intensity will be expressed as quantity per unit of length, of area or of volume. Strictly speaking, the intensity of a quantity at a point P on a range is the limit of the ratio of ΔQ to ΔR, where ΔQ is that part of the quantity Q which pertains to the part ΔR of the range R, and the limit is taken by allowing ΔR to approach zero, always including the point P.

If the differential element of the range is expressed as $f(x)dx$, and if the intensity of the distributed quantity pertaining to this element may be expressed as $\phi(x)$, then the total amount of the quantity between x_1 and x_2 is

$$Q = \int_{x_1}^{x_2} \phi(x)\, f(x)\, dx,$$

and if we divide this result by the range,

$$R = \int_{x_1}^{x_2} f(x)\, dx,$$

we obtain the average intensity of Q on R. Similar results may be stated, whatever form of differential is used.

The applications of this general idea are very important, and extend to many fields of engineering and physical science. We give here a few of the more direct applications.

1. **Distributed masses.** If mass is distributed over a range, the intensity of mass is called **density.** The terms line density, surface density, volume density, are used, according to the character of the

range. Consider a triangle (Fig. 74) with base b and altitude h, with base on the x-axis as shown. Let the density be a function, of x, say $\sin \frac{\pi x}{b}$, then corresponding to the differential element of area $dA = \frac{h}{b} x\, dx$, we have the mass

$$dm = \frac{h}{b} x \sin \frac{\pi x}{b}\, dx,$$

and the total mass pertaining to the triangle is

$$m = \int_0^b \frac{h}{b} x \sin \frac{\pi x}{b}\, dx.$$

Here the density is surface density, and the form obtained may very well pertain to a triangular plate of varying thickness, but of homogene-

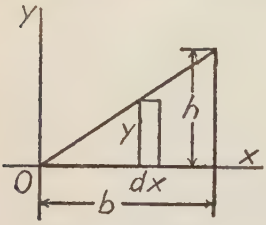


Fig. 74

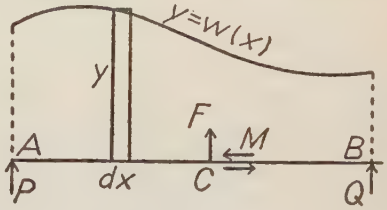


Fig. 75

ous material. The first and second moments of this mass with respect to the y-axis are given by the integrals.

$$M_y = \int_0^b \frac{h}{b} x^2 \sin \frac{\pi x}{b}\, dx,$$

$$I_y = \int_0^b \frac{h}{b} x^3 \sin \frac{\pi x}{b}\, dx.$$

If we divide M_y by the mass m we obtain the abscissa of the center of gravity of the mass, and if we divide I_y by m we obtain the squared radius of gyration with respect to the y-axis. If the density were constant over the area, the integrals here would reduce to this constant times the integrals for area and moments as considered in the previous article. If now we wish to find moments of this mass with respect to the x-axis, the density being a function of x makes it desirable to retain the same form for the element of mass, and by the

previous article the mean value of y on this element is $\frac{hx}{2b}$, and the mean value of y^2 is $\frac{h^2x^2}{3b^2}$, so that we may write

$$M_x = \frac{h}{2b} M_y, \qquad I_x = \frac{h^2}{3b^2} I_y.$$

Consideration of this example should make sufficiently clear the methods of finding masses and their moments, whatever the form of differential of range, provided only that the density may be expressed as a function of the variable of integration.

2. **Forces on moving bodies.** If a body of mass m is acted upon by a force F in a fixed direction, the motion of the body in this direction is accelerated according to the law

$$m \frac{d^2s}{dt^2} = F. \tag{1}$$

where s is displacement in the given direction and t is time, the second derivative being acceleration. If F is expressed as a function of the time, and if $v = \frac{ds}{dt}$, we may write

$$m \frac{dv}{dt} = F,$$

or, if $v = v_0$ when $t = t_0$ and $v = v_1$ when $t = t_1$

$$m(v_1 - v_0) = \int_{t_0}^{t_1} F \, dt. \tag{2}$$

The integral on the right is called the **impulse** of the force, and we thus see that force may be regarded as the intensity of an impulse distributed throughout a time interval. The product of a mass by its velocity is called its **momentum.** The equation (2) states that the change of momentum is equal to the impulse on the body.

As the mass m moves, the force F, changing from time to time, takes different values at different positions of the body and may, therefore, be regarded as a function of s. Equation (1) may be written

$$m \frac{ds}{dt} \frac{d^2s}{dt^2} = F \frac{ds}{dt},$$

or

$$m\, v\, dv = F\, d\, s,$$

and integrating as before

$$m\left(\frac{v_1^2}{2} - \frac{v_o^2}{2}\right) = \int_{s_o}^{s_1} F\, ds. \tag{3}$$

The integral on the right is called the **work** of the force, and it is clear that force may be regarded as the intensity of work distributed along the path of motion. The quantity $\frac{1}{2}\, mv^2$ is called the **kinetic energy** of the mass, and equation (3) states that the change of kinetic energy is equal to the work done on the body. Let the work done be denoted by U, then

$$\frac{dU}{ds} = F.$$

and therefore,

$$\frac{dU}{dt} = \frac{dU}{ds} \cdot \frac{ds}{dt} = Fv.$$

This quantity, the rate of doing work, is called **power.** Notice that force is the intensity of work distributed through the distance moved, while power is the intensity of work distributed through the time interval. If the power P is regarded as a function of the time, then

$$U = \frac{m}{2}\,(v_1^2 - v_o^2) = \int_{s_o}^{s_1} F\, ds = \int_{t_o}^{t_1} P\, dt. \tag{4}$$

3. **Distributed forces.** We shall consider the methods of finding the effects of forces that are distributed as to place of application. Our objects are best attained by indicating the treatment of problems of certain specific types.

(a) **The loaded beam.** Consider a beam AB of length L (Fig. 75), supported at the ends by vertical forces P and Q, and sustaining a load whose intensity varies from point to point along the beam. Take the origin at A, and let x be measured along the beam AB. The weight of load per unit of length at distance x from A is a function of x, say w(x). The graph of this function is a curve whose ordinate is the intensity of the load. The total load is

$$P + Q = \int_o^L w(x)\, dx, \tag{5}$$

and the moment of the load with respect to A is

$$QL = \int_0^L x\, w(x)\, dx. \tag{6}$$

The function w(x) being known, these equations are sufficient to determine the supporting forces P and Q.

If at any point, C, of the beam we consider a plane cross section separating the beam into two parts, the two parts react on each other by forces called **internal stresses.** The vertical component of the force with which the part AC acts on the part CB is called **shearing force.** Denote this by F, then AC is acted upon at C by a force $-F$. Since the vertical forces on AC must balance we have, putting the distance AC equal to a,

$$F = P - \int_0^a w(x)\, dx. \tag{7}$$

Now the resultant of these vertical forces on AC is at most a couple, and must be balanced by a couple M at C, called the **bending moment.** The moment of a couple is the same with respect to all points, therefore, taking moments about A,

$$M = aF + \int_0^a x\, w(x)\, dx = aP - a\int_0^a w(x)\, dx + \int_0^a x\, w(x)\, dx. \tag{8}$$

By performing the integrations in (7) and (8) we may express F and M as functions of a, that is, as functions of the abscissa of C.

(b) **Water pressure.** We find in hydro-mechanics that the force exerted by a liquid on a surface is always normal to the surface. The intensity of force on a surface, that is, force per unit area, is called **pressure.** It is shown that the difference of pressure between two points of a liquid is equal to the weight per unit volume of the liquid multiplied by the difference of altitude between the points, regardless of the inclination of the surfaces. The weight of a cubic foot of fresh water is about 62.5 pounds, therefore the pressure at depth h feet below a free surface of water is 62.5 h pounds per square foot above atmospheric pressure.

Consider now any plane area inclined at an angle ϕ to the vertical and let the horizontal width at depth y below the free surface be denoted by x, the element of area is, then, $x \sec \phi\, dy$, and

$$A = \sec \phi \int_{y_1}^{y_2} x\, dy. \tag{9}$$

where y_1 and y_2 are the extreme values of y on the area. The total pressure is

$$P = w \sec \phi \int_{y_1}^{y_2} y\, x\, dy, \tag{10}$$

where w is weight per unit volume of the liquid. The moment of this pressure with respect to the line in which the plane of the area intersects the free surface of the liquid may be written

$$P\, y_c \sec \phi = w \sec^2 \phi \int_{y_1}^{y_2} y^2\, x\, dy, \tag{11}$$

since the distance from this line to an element of area at depth y is $y \sec \phi$. This equation defines the quantity y_c, which is called the **depth of the center of pressure.** If the area in question lies in a vertical plane we have $\sec \phi = 1$, and in this case the three integrals represent the area and its first and second moments with respect to the line at the free surface.

(c) **Stresses.** Let the area in Figure 76 represent a cross section of a beam, column or strut. Suppose the area symmetrical with respect to the line CD, and let AB be perpendicular to CD at O, the center of gravity of the area. This plane section divides the body into two parts which act on each other by forces distributed over this surface. The external forces acting on one of the two parts must be balanced by those acting on the other part, and the reaction at this section must transmit these forces from one part to the other. The intensity of force normal to the plane section at any point is called **normal stress,** while the intensity of stress in the plane section is called **tangential** or **shearing stress.**

Let us suppose that the resultant of forces acting on one part of the body is a shearing force F acting along the line CD, a thrust P acting normal to the plane at O, and bending moment M which is a couple (see Fig. 75) having moment M with respect to AB. At present we confine attention to the distribution of normal stress over the plane section, which must be such as to give a total force equal to P and a moment with respect to AB equal to M.

In the mechanics of materials assumptions are made which yield the conclusion that the normal stress is a linear function of the distance y from AB. Denote this by

$$p = k(y - b), \tag{12}$$

then denoting the element of area by x dy we have for the total normal force

$$P = k \int_{y_1}^{y_2} (y - b)\, x\, dy, \tag{13}$$

and for the moment with respect to AB we have

$$M = k \int_{y_1}^{y_2} y(y - b)\, x\, dy. \tag{14}$$

Having performed these integrations we may solve for k and b in terms of P and M, then equation (12) gives us the stress at distance y from AB. The extreme values of p are at $y = y_1$ and $y = y_2$. The

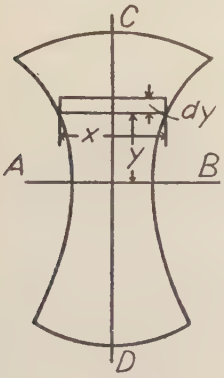


Fig. 76

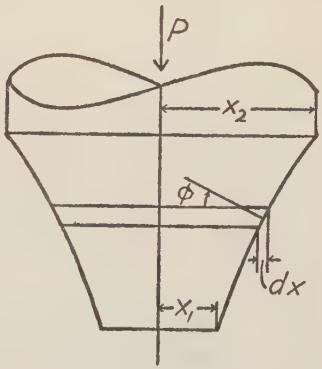


Fig. 77

line y = b at which p = σ is called the **neutral axis** of the section. The integral in (13) is the first moment of the area with respect to the line y = b, and if P = o this moment is zero, so that b = o, and the neutral axis becomes the line AB. In this case the moment M is proportional to the second moment of the area with respect to AB.

(d) **Pivot friction.** Consider a pivot bearing in the form of a surface of revolution, in which a shaft turns subject to the thrust P (Fig. 77). There is a normal pressure on the surface of the bearing, which we will assume to be a constant, p. Let ϕ be the angle that the normal to the surface makes with the horizontal, then the element of area may be written 2π x csc ϕ dx and the vertical component of pressure on this element is p sin ϕ times this, therefore

$$P = 2\pi p \int_{x_1}^{x_2} x\, dx = p(\pi x^2 - \pi x_1^2). \tag{15}$$

This shows that the pressure p is equal to the thrust P divided by the vertical projection of the bearing area.

If the coefficient of friction in the bearing is μ, then μp is the force per unit of area which resists turning of the shaft. The total moment of this resistance about the shaft axis is

$$T = 2\pi\mu p \int_{x_1}^{x_2} x^2 \csc\phi \, dx, \tag{16}$$

where, of course, $\csc\phi \, dx = ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \, dx$. In the case of a conical pivot ϕ is constant, and if $\phi = \frac{\pi}{2}$ the bearing becomes a flat collar or step bearing.

EXERCISES XL

1. Find the center of gravity of a semi-circular plate whose thickness at any point is proportional to the distance of the point from the bounding diameter. (Take dA = 2ydx, and surface density as kx.)

2. Find the moment of inertia of the plate in 1 with respect to the diameter.

3. Find the center of gravity of a rectangular plate whose thickness is proportional to the square of the distance from one edge. Express your result in terms of the second and third moments of the area with respect to the sharp edge.

4. A body weighing 100 pounds is acted upon by a force $F = 2t + 3e^{-t}$ where F is in pounds and t in seconds. What change of velocity occurs from t = o to t = 2?

5. The stroke of a certain piston is three feet. For the first foot the total steam pressure on the piston is F = 100,000 pounds. For the remainder of the stroke the pressure changes according to the law Fx = 100,000, where x is distance in feet that the piston moves. How much work is done during the stroke by the steam on one side of the piston?

6. The force acting on a body weighing 320 pounds is F = 100 t, where F is in pounds and t in seconds. If v = o when t = o, find v as a function of t. Also find the power as a function of t, and find the work done in three seconds.

7. Taking the equation for v in 6, find s as a function of t, then express F as a function of s. Compute the work of the force during three seconds by the relation $U = \int F ds$, and compare with the final result in 6.

8. A beam carries a load whose intensity in pounds per foot run is $w = 1{,}000\ x$ where x is distance in feet from one end of the beam. If the beam is 10 feet long, find the supporting forces at the ends and find the shearing force and bending moment at any distance, x, from the end.

9. Solve 8 after changing so that $w = 500 + 1{,}000 \sin \frac{\pi x}{10}$.

10. The end of a tank has the form of a trapezoid, upper base 16 feet, lower base 8 feet, altitude 12 feet. If water stands 9 feet deep in the tank, how high is the center of pressure on the end of the tank.

11. An automatic water gate 4 feet high by 5 feet wide is hinged at the upper edge and rests against a stop at the lower edge. Find the force on the hinges if water stands three feet high on one side and one foot high on the other.

12. A cylindrical container is inclined at an angle of 45°, and contains liquid just sufficient to submerge the circular base. Find the center of pressure on this base.

13. Find the greatest normal stress in the section of a beam 2 inches wide by 8 inches deep, where the bending moment is 5 tons-inches.

14. If the beam in 13 is subject to a pull of 4 tons in addition to the bending moment, what is the greatest normal stress, and where is the neutral axis?

15. If a circular shaft of radius r is subject to a bending moment M what is the maximum normal stress? If the shaft is hollow, inner radius r_1, what does the stress become?

16. The bearing surface of a pivot is a cone, radius 3 inches and altitude 4 inches. The thrust of the shaft is 1500 pounds, the coefficient of friction is $\frac{1}{20}$. Find the moment of resistance of friction.

17. Find the moment of the friction on a collar bearing 4 inches outside radius and 3 inches inside radius, when subjected to a total pressure of 4000 pounds, assuming $\mu = 0.03$.

Art. 44. Successive Integrations

1. **Repeated integration in one variable.** The indefinite integral of a function is itself a function and may be integrated again. Thus, we have the notation

$$\int \left[\int f(x)\, dx \right] dx = \int\int f(x)\, dx^2$$

to indicate the function whose second derivative is f(x). Similarly the function whose third derivative is f(x) is indicated by

$$\int\int\int f(x)\, dx^3,$$

and so on to any order. These functions contain one arbitrary constant for each integration indicated, and are completely determined only when the function f(x) is known and initial conditions are given sufficient to determine the constants.

For example, suppose the acceleration is given as a function of the time

$$\frac{d^2s}{dt^2} = \sin \pi t$$

and, with the understanding that $\frac{ds}{dt} = 0$ and $s = 2$ when $t = o$, let it be required to find s when $t = 5$. We have

$$\frac{ds}{dt} = -\frac{1}{\pi}\cos \pi t + c_1, \quad c_1 = \frac{1}{\pi},$$

$$s = -\frac{1}{\pi^2}\sin \pi t + \frac{t}{\pi} + c_2, \quad c_2 = 2.$$

And when $t = 5$ we have $s = 2 - \frac{5}{\pi}$. This result may be indicated by the single expression

$$s_5 = 2 + \int_o^5 \int_o^t \sin \pi t\, dt^2,$$

and if we desire to indicate s as a function of t we have to put t for 5 in this expression.

We have seen that an integral of a function may be interpreted as an area. Similarly the second integral may be interpreted as the first moment of an area and the third integral as the second moment of an area. The area under the curve $y = f(x)$ from a to x (Fig. 78) is

$$A = \int_a^x f(x)\, dx, \tag{1}$$

and if we denote the first moment with respect to the ordinate at x of the area A by M we have

$$M + \Delta M = M + A\, \Delta x + \Delta A \cdot c\, \Delta x$$

where c is approximately one-half. We find

$$\frac{\Delta M}{\Delta x} = A + c\,\Delta A, \qquad \frac{dM}{dx} = A,$$

therefore

$$M = \int_a^x \int_a^x f(x)\,dx^2. \tag{2}$$

By similar reasoning it may be shown that the second moment of A with respect to the ordinate at x is

$$I = 2\int_a^x \int_a^x \int_a^x f(x)\,dx^3. \tag{3}$$

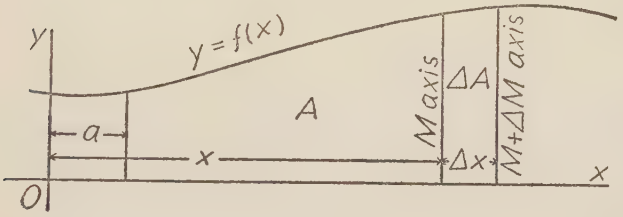


Fig. 78

In (2) and (3) the successive integrations are from the same initial point, x = a, and the result becomes a definite integral when the upper limit for the last integration is fixed. If the initial points are different for successive integrations, or if upper limits other than the last are different from x, the result is still capable of interpretation in terms of areas and moments, but with less facility and profit.

A direct application of (1) and (2) occurs in the theory of beams. Let a beam lie along Ox (Fig. 78), and let f(x) be the intensity of force downward on the beam at any point x, then the area A is the total load between a and x. Let the shearing force and bending moment at a be F_a and M_a, respectively. The shearing force at x is

$$F_x = F_a - A = F_a - \int_a^x f(x)\,dx. \tag{4}$$

The bending moment at x consists of three parts. M_a, being a couple, acts with the same effect at x as at a, F_a has a moment $F_a x$ at x, and the distributed load has a moment at x given by (2).

Therefore

$$M_x = M_a + F_a x - \int_a^x \int_a^x f(x)\, dx^2 = M_a + \int_a^x F_x dx. \qquad (5)$$

It should be understood that a concentrated load or support for the beam between a and x amounts to an infinite discontinuity in f(x), and gives a finite discontinuity in F_x but leaves M_x continuous. This problem, which cannot be discussed completely here, is handled by treating separately those parts of the beam over which f(x) is continuous, and adjusting constants of integration to satisfy the given conditions. The equations (4) and (5) are fundamental in the theory of beams under distributed loads.

2. **Iterated integrals in two variables.** In previous articles the differential of a quantity has always been expressed as a differential of an independent variable multiplied by a function of that variable, and in using such expressions as ydx or $\sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$ for the differential of area or of arc we have understood the coefficient of dx to be a function of x, and we have always so expressed it before integrating. Moreover, in finding moments and in dealing with distributed quantities the function by which we have multiplied the differential of arc, area or volume has always been expressed in terms of this variable.

Now in the case of the differential of arc this is no restriction, since the equation of the curve whose arc is considered expresses one variable in terms of the other, but in the case of the differential of area or volume the restriction amounts to a serious limitation. We may now show how to avoid this restriction in the case of a differential area.

First let us return to the subject of moments treated in Art. 42, and consider again the area under the curve $y = \sin x$ from $x = 0$ to $x = \pi$. We have $dA = \sin x\, dx$, and to find the first moment with respect to the x axis we observe that the mean value of y on a rectangle whose height is $\sin x$ is $\bar{y} = \frac{1}{2} \sin x$ and write

$$M_x = Ay_m = \int_0^\pi \frac{\sin^2 x}{2}\, dx = \frac{\pi}{4}.$$

Now this mean value of y is known as the result of a previous integration,

$$\bar{y} \sin x\, dx = dx \int_0^{\sin x} y\, dy = \frac{1}{2} \sin^2 x\, dx.$$

Here dx is the constant width of the rectangle, and may be divided out of the equation. We may, therefore, write

$$M_x = Ay_m = \int_0^{\pi}\left[\int_0^{\sin x} y\,dy\right]dx = \int_0^{\pi}\int_0^{\sin x} y\,dy\,dx.$$

Similarly the second moment of this area with respect to the x-axis is given by

$$I_x = A(y^2)_m = \int_0^{\pi}\int_0^{\sin x} y^2\,dy\,dx.$$

In general the expression

$$\int_{x_1}^{x_2}\int_{f_1(x)}^{f_2(x)} \phi(x, y)\,dy\,dx,$$

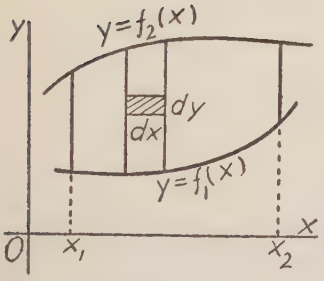


Fig. 79

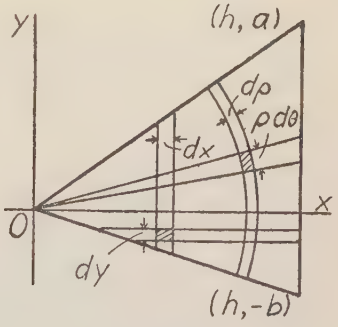


Fig. 80

is known as an **iterated integral,** and if $\phi(x, y)$ is regarded as giving the intensity of a quantity at the point (x, y) the integral is the total of this quantity on the area bounded by th e curves $y = f_1(x)$ and $y = f_2(x)$, and the lines $x = x_1$ and $x = x_2$ (Fig. 79).

As an illustration let us find the polar moment of inertia of a triangle with respect to a vertex. We shall take the origin at this vertex and the x-axis perpendicular to the opposite side (Fig. 80). Let the coordinates of the other two vertices be (h, a) and (h, $-$b), then the bounding lines are $y = \frac{a}{h}x$, $y = \frac{-b}{h}x$ and $x = h$.

The squared distance from O to any point (x, y) is $x^2 + y^2$ and this

may be regarded as the "intensity" of polar moment of inertia at that point. The total is

$$I_o = \int_o^h \int_{\frac{-b}{h}x}^{\frac{a}{h}x} (x^2 + y^2)\, dy\, dx.$$

Integrating first with respect to y, regarding x as a constant, that is, finding the total for a vertical strip of width dx, we have

$$I_o = \int_o^h \left[x^2 y + \frac{y^3}{3} \right]_{\frac{-b}{h}x}^{\frac{a}{h}x} dx = \int_o^h \left[\frac{(a+b)}{h} + \frac{a^3 + b^3}{3h^3} \right] x^3 dx$$

and now integrating with respect to x

$$I_o = \left[\frac{3h^2(a+b) + (a^3+b^3)}{12h^3}\, x^4 \right]_o^h = \frac{3h^3(a+b) + h(a^3+b^3)}{12}.$$

For further illustration we shall now solve the same problem by integrating with respect to x first. Now the limits for x will be functions of y, and will be differently expressed when y is between $-b$ and o and between o and a. We write

$$I_o = \int_{-b}^{o} \int_{-\frac{h}{b}y}^{h} (x^2 + y^2)\, dx\, dy + \int_o^a \int_{\frac{h}{a}y}^{h} (x^2 + y^2)\, dx\, dy$$

$$= \int_{-b}^{a} \left[\frac{h^3}{3} + hy^2 + y^3 \left(\frac{h^3}{3b^3} + \frac{h}{b} \right) \right] dy + \int_o^a \left[\frac{h^3}{3} + y^2 h - y^3 \left(\frac{h^3}{3a^3} + \frac{h}{a} \right) \right] dy$$

$$= \frac{3h^3 b + hb^3}{12} + \frac{3h^3 a + ha^3}{12} = \frac{3h^3(a+b) + h(a^3+b^3)}{12}.$$

The method is available also in polar coordinates in which case the element of area takes the form $\rho d\theta d\rho$ (see fig. 80). If we integrate this with respect to ρ from ρ_1 to ρ_2 we have $\frac{1}{2}(\rho^2{}_2 - \rho^2{}_1)d\theta$, and if we integrate with respect to θ from θ_1 to θ_2 we have $\rho(\theta_2 - \theta_1)d\rho$. The limits ρ_1 and ρ_2 may be functions of θ, and the limits θ_1 and θ_2 may be functions of ρ, and these are the two forms of differential of area in polar coordinates used in Art. 41.

To illustrate the iterated integral in polar coordinates let us solve again the problem on the triangle (Fig. 80). The squared distance is ρ^2, and the bounding lines are $\theta = \tan^{-1}\left(\frac{-b}{h}\right)$, $\theta = \tan^{-1}\frac{a}{h}$, $\rho = h \sec \theta$.

We write

$$I_o = \int_{\tan^{-1}\left(\frac{-b}{h}\right)}^{\tan^{-1}\frac{a}{h}} \int_o^{h\sec\theta} \rho^3 d\rho\, d\theta$$

$$= \int_{\tan^{-1}\left(\frac{-b}{h}\right)}^{\tan^{-1}\frac{a}{h}} \frac{h^4 \sec^4\theta}{4}\, d\theta$$

$$= \frac{h^4}{4}\left[\tan\theta + \frac{\tan^3\theta}{3}\right]_{\tan^{-1}\left(\frac{-b}{h}\right)}^{\tan^{-1}\frac{a}{h}}$$

$$= \frac{3h^3(a+b) + h(a^3+b^3)}{12}.$$

If we change the order of integration we must break the range for ρ into three parts, using different limits for θ on each part.

$$I_o = \int_o^h \int_{\tan^{-1}\left(\frac{-b}{h}\right)}^{\tan^{-1}\frac{a}{h}} \rho^3\, d\,\theta\, d\rho$$

$$+ \int_h^{\sqrt{h^2+b^2}} \int_{\tan^{-1}\left(\frac{-b}{h}\right)}^{-\sec^{-1}\frac{\rho}{h}} \rho^3\, d\theta\, d\rho$$

$$+ \int_h^{\sqrt{h^2+a^2}} \int_{\sec^{-1}\frac{\rho}{h}}^{\tan^{-1}\frac{a}{h}} \rho^3\, d\theta\, d\rho.$$

Quite obviously, this order is less convenient than the previous one, and we shall not take the space to complete the integration.

It will be evident that four methods are available for applying iterated integrals to any problem dealing with moments or distributed quantities on an area, and that the choice of method should be governed by consideration both of the form of the function which shows the distribution or intensity and of the character of the bounding curves. The advantage of the iterated integral is that it permits the integrand to be a function of either or both variables.

EXERCISES XLI

1. Given $s = 5 + \int_0^t \int_2^t 12dt^2$, find s as a function of t, and find velocity and acceleration at $t = 0$. For what value of t is $v = 0$.

2. If $\frac{d^2s}{dt^2} = -24t$, and $v = 0$ when $t = 5$ and $s = 0$ when $t = 3$, express s in terms of t by the symbol of integration without integrating. Check your result by evaluating your integral, then finding velocity and acceleration by differentiation.

3. Apply equations (4) and (5) in section 1 of the present article to a beam of length L, supported at the ends and loaded uniformly, taking M_a and F_a at the origin, the left end. Check by use of equations in section 3, Art. 43.

4. Apply equations (2) and (3) in section 1 to a rectangle, and check by the method of Art. 42.

5. Apply the equations and the check as in 4 to the triangle bounded by the lines $y = 3x$, $y = 0$, $x = a$.

6. In equations (2) and (3) of section 1 take $f(x) = \sin x$ and apply the check from Art. 42.

7. The moment of inertia with respect to the y-axis of a certain area is

$$I_y = A(x^2)_m = \int_0^5 \int_{y+1}^{6-y} x^2 \, dx \, dy,$$

draw the area and find $(x^2)_m$.

8. To what area does the following integral apply?

$$\int_2^5 \int_{-x}^{x^2} x \, y \, dy \, dx.$$

This integral, the element of area dy dx being multiplied by the product xy, is called the **product of inertia** of the area with respect to the coordinate axes.

9. Find the product of inertia of a rectangle with respect to two edges. (See 8.)

10. Find the product of inertia of a quadrant of a circle with respect to the bounding diameters. (See 8.)

11. Solve problem 10 in polar coordinates, observing that $xy = \rho^2 \sin \theta \cos \theta$.

12. Show that the polar moment of inertia of a rectangle with respect to a corner is equal to the sum of the moments of inertia with respect to two edges.

13. Find the mean value of $\sqrt{\frac{x}{y}}$ on the area bounded by $x = y^2$ and $y = x^2$.

14. Find the mean value of $\sec \theta$ on the area given by the following integral, and draw the area.

$$\int_{-\frac{\pi}{6}}^{\frac{\pi}{6}} \int_{0}^{\cos\theta} \rho \, d\rho \, d\theta.$$

Art 45. Substitutes for Integration

As we have seen, the integral of a function between limits may always be interpreted as an area. If, then, we are able by any means whatever to find the area in question we shall have the value of the integral. This fact is of extreme importance when the integrand is of such character that the integration cannot be performed, or when the function is known only as a graph or by tabulated values. It is the purpose of the present article to present certain standard methods of computing areas from the tabulated values, and to show their application to problems involving integration.

1. **Numerical approximation of a definite integral.** Let us consider the integral

$$A = \int_{x_0}^{x_n} f(x)\, dx.$$

and let us suppose that the ordinates of the curve $y = f(x)$ are known at x_0, $x_0 + h$, $x_0 + 2h$ . . . . $x_0 + \frac{n}{2} h$. Let these ordinates be represented by y_0, y_2, y_4 . . . . y_n, repectively, shown by solid lines in the figure (Fig. 81). The area between two of these ordinates is approximately a trapezoid, and the whole area is approximately the sum of such trapezoids.

$$A_t = \frac{h}{2}\Big[(y_0 + y_2) + (y_2 + y_4) + \ldots + (y_{n-2} + y_n)\Big] \qquad (1)$$

$$= \frac{h}{2}\Big[(y_0 + y_n) + 2(y_2 + y_4 + \ldots + y_{n-2})\Big].$$

The use of even subscripts here is a convenience in relation to what follows, and is not essential to the formula. If each subscript is divided by 2 we have, in its usual form, what is known as the **trapezoidal rule.**

If we consider only the ordinates at the mid-points of each of the equal intervals, denoted by $y_1, y_3, \ldots . y_{n-1}$ and shown as broken lines, then the expressions hy_1, hy_3, etc., are approximately equal to the areas on the respective intervals, and an approximation for the whole area is

$$A_m = h(y_1 + y_3 + \ldots . + y_{n-1}). \qquad (2)$$

This is known as the **mid-ordinate rule,** and of course the use of odd subscripts only is not an essential feature of the rule.

It may be observed that the area on an interval where the curve is concave downward, as from y_2 to y_4, is underestimated in the trapezoidal rule and overestimated in the mid-ordinate rule, while the situation is reversed where the curve is concave upward, as from y_6 to y_8. If an area is computed by both methods and the two estimates A_t and

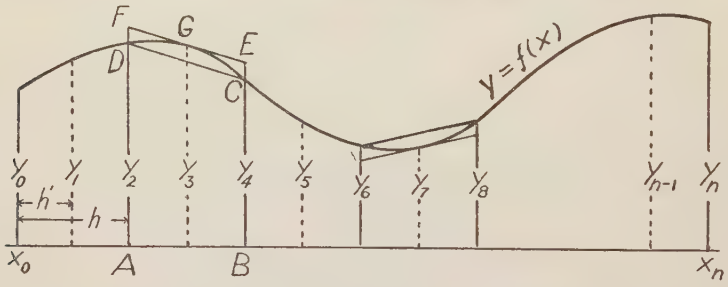


Fig. 81

A_m, are materially different then the true area is almost certainly between these two estimates. If we take their mean we have, putting $h = 2h'$.

$$\frac{1}{2}(A_t + A_m) = \frac{h'}{2}\Big[(y_0 + y_n) + 2(y_1 + y_2 + \ldots . y_{n-1})\Big].$$

and this is clearly a statement of the trapezoidal rule, using both odd and even ordinates and having the equal intervals of width h'.

If we consider a given interval of width h, as AB, we see that the trapezoidal rule uses the area ABCD, while the mid-ordinate rule uses ABEF, and on account of the curvature of DGC it is apparent that the latter is the more accurate. To obtain an estimate of the area DCG we may consider a simple curve passing through these three points, and the simplest curve for the purpose is a parabola of the type $y = ax^2 + bx + x$. It is easily shown, then, that EF, parallel to DC, is

tangent to the parabola at G and that the area DCG is two-thirds of the area of the parallelogram DCEF. Taking the parabola through the three points as an approximation to the curve on this interval we should have the relation

$$\text{ABCGD} = \text{ABCD} + \frac{2}{3}\,\text{DCEF}$$

$$= \text{ABCD} + \frac{2}{3}(\text{ABEF} - \text{ABCD}).$$

Since this method of estimating the area is applicable to every interval of width h, we have for an approximation to the whole area

$$A_s = A_t + \frac{2}{3}(A_m - A_t) = \frac{1}{3}(A_t + 2A_m)$$

$$= \frac{h'}{3}[(y_0 + y_n) + 4(y_1 + y_3 + \ . \ . \ . \ . \ + y_{n-1}) \quad (3)$$

$$+ 2(y_2 + y_4) + \ . \ . \ . \ . \ + y_{n-2})].$$

This is known as **Simpson's rule,** but to distinguish it from other rules given by the same author it is variously designated as Simpson's parabolic rule, Simpson's one-third rule, and Simpson's first rule.

We note that Simpson's rule requires that an odd number of equally spaced ordinates be known, while the two previous rules require only that the ordinates be equally spaced. When the data is in proper form for its use, Simpson's rule usually gives far more accurate results than either of the other rules. In case the ordinates tabulated are not equally spaced, the principle involved in the trapezoidal rule will be most convenient in estimating the area, each trapezoid being treated separately and the parts added to obtain the whole area. It may, in certain cases, be profitable to treat portions of the area by different rules, according to the nature of the data available.

The error in computing the area by any one of these rules depends on the magnitude of h, and if we may take h sufficiently small the error may be made as small as desired. For a given value of h, Simpson's rule will usually be found more accurate than the either of the other rules would be with an h half as large, that is, with double the number of divisions.

2. **Construction of the integral curve.** If we have a curve $y = f(x)$ given by a graph or by tabulated values we may make use of the interpretation of integral as area to construct an integral curve. Of course, on account of the constant of integration, the integral curve may go through any desired point, and if we take this point as (x_o, A_o), we have for the ordinate A_x of the integral curve at any x,

$$A_x = A_o + \int_o^x f(x)\, dx.$$

Now, if the ordinates of the curve $y = f(x)$ at x_o, x_1, x_2, etc., are known as y_o, y_1, y_2, etc. (Fig. 81), then the value of A_x at x_1 is approximately

$$A_1 = A_o + \frac{1}{2} h_1 (y_o + y_1),$$

where $h_1 = x_1 - x_o$. And similarly at any point x_n,

$$A_n = A_{n-1} + \frac{1}{2} h_n (y_{n-1} + y_n). \tag{4}$$

where $h_n = x_n - x_{n-1}$. This method corresponds to the trapezoidal rule, and it does not require that the ordinates be equally spaced. We may apply a method corresponding to the mid-ordinate rule by putting in place of $\frac{1}{2}(y_{n-1} + y_n)$ in equation (4) the ordinate at the mid-point of the interval from x_{n-1} to x_n.

If the ordinates are equally spaced we may apply Simpson's rule to give a more accurate approximation,

$$A_n = A_{n-2} + \frac{h}{3} (y_{n-2} + 4y_{n-1} + y_n), \tag{5}$$

where h is the common distance between ordinates. This formula does not apply at $n = 1$, unless y_{-1} at a distance h to the left of x_o is known. But A_1 may be found from (4) in any case.

3. **Application to problems in integration.** Let us consider the area under a curve $y = f(x)$, and let us suppose the data given are contained in the columns of the tables headed x, y (even) and y (odd).

x	y		xy		y^2	
	(Even)	(Odd)	(Even)	(Odd)	(Even)	(Odd)
0	(1.2)		(o)		(1.44)	
0.4		3.8		1.52		14.44
0.8	4.6		3.68		21.16	
1.2		4.9		5.88		24.01
1.6	4.2		6.72		17.64	
2.0		2.8		5.60		9.64
2.4	(1.0)		(2.40)		(1.00)	
	(2.2)+ 8.8	11.5	(2.4)+ 10.40	13.00	(2.44)+ 38.80	48.09

Here h = 0.4, and by the trapezoidal rule we should have for the area from x = 0 to x = 2.4,

$$A_t = \frac{0.4}{2}[2.2 + 2(8.8 + 11.5)] = 8.56,$$

while by Simpson's rule we have

$$A_s = \frac{0.4}{3}[2.2 + 2(8.8) + 4(11.5)] = 8.77.$$

These are approximations for the integral

$$A = \int_0^{2.4} f(x)\,dx.$$

If we wish to find the center of gravity of this area we must approximate the integrals

$$Ax_m = \int_0^{2.4} xy\,dx, \quad Ay_m = \frac{1}{2}\int_0^{2.4} y^2\,dx.$$

Accordingly we form the columns for xy and y^2, keeping separate columns for the odd ordinates for convenience in separate addition for Simpson's rule. By the trapezoidal rule we have

$$A_t x_m = \frac{0.4}{2}[2.4 + 2(23.40)] = 9.84,$$

$$A_t y_m = \frac{1}{2}\cdot\frac{0.4}{2}[2.44 + 2(86.89)] = 17.62,$$

from which $x_m = 1.15$ and $y_m = 2.06$. But by Simpson's rule we have

$$A_s x_m = \frac{0.4}{3}[2.4 + 2(10.40) + 4(13.00)] = 10.00,$$

$$A_s y_m = \frac{1}{2} \cdot \frac{0.4}{3}[2.44 + 2(38.80) + 4(48.09)] = 18.16.$$

and therefore $x_m = 1.14$ and $y_m = 2.07$.

To find the moments of inertia with respect to the axes of reference we should write

$$A(x^2)_m = \int_0^{2.4} x^2 y \, dx, \; A(y^2)_m = \frac{1}{3}\int_0^{2.4} y^3 \, dx,$$

and proceed to evaluate these integrals by forming the columns for x^2y and y^3. If we wish to find the product of inertia we write

$$\int_0^{2.4}\int_0^{y} x\, y \, dy \, dx = \frac{1}{2}\int_0^{2.4} x\, y^2 \, dx,$$

and evaluate by forming the column for xy^2. To find the polar moment of inertia with respect to the origin we have

$$\int_0^{2.4}\int_0^{y} (x^2 + y^2) \, dy \, dx = \int_0^{2.4} x^2 y \, dx + \frac{1}{3}\int_0^{2.4} y^3 \, dx,$$

and this is found to be the sum of the two moments of inertia indicated above.

To find the volume of revolution formed by revolving this area about the x-axis we have

$$V_x = \pi \int_0^{2.4} y^2 \, dx = 2\pi A y_m,$$

and the last expression is known from the computation above. If the area were revolved about the y-axis each element of area $y\, dx$ would generate the volume $2\pi\, x\, y\, dx$, hence the whole volume is

$$V_y = 2\pi \int_0^{2.4} x\, y \, dx = 2\pi A x_m,$$

and may be found from a previous computation.

The integral

$$V = \int_{x_0}^{x_n} \phi.(x) \, dx$$

represents a volume of which the area of a section perpendicular to the x-axis at any x is $\phi(x)$. If this cross-sectional area is known at equal intervals the value V may be found by one of our three rules. In applying Simpson's rule we should have for the area from x_{n-2} to x_n

$$\frac{h}{3}[\phi(x_{n-2}) + 4\phi(x_{n-1}) + \phi(x_n)]$$

and this is identical with the **prismoid rule** for finding the volume of a solid between parallel bases in terms of the areas of the two bases and that of the mid-section. The formula gives exact results for a great variety of forms and is a satisfactory approximation under conditions similar to those indicated for areas.

EXERCISES XLII

1. Using a table of natural sines, and taking $h = \frac{\pi}{6}$, apply the trapezoidal rule to find the area under the sine curve from o to π, and the first and second moments of this area with respect to the axes.

2. Apply the mid-ordinate rule to problem 1, using the same h, and compare results with the result of actual integration.

3. Treat the foregoing problem by Simpson's rule, taking $h = \frac{\pi}{4}$, and again taking $h = \frac{\pi}{6}$, and compare results with the mean of the results by the two previous methods.

4. Corresponding values of x and y are as follows:

x.....	0	0.8	1.9	3.1	4.3	5.5	6.6	7.9	9.0	10.0
y.....	2.00	2.10	2.32	2.69	3.11	3.69	4.30	5.17	6.00	6.91

Using cross-section paper, plot the points and draw a smooth curve through these points. Determine the ordinates of this curve at integral values of x, and using these ordinates determine, by an application of Simpson's rule, the area under the curve, the volume of revolution about the x-axis, and the center of gravity of the area.

5. Write the values of e^{-x^2} at intervals of 0.1 from x = 0 to x = 1 from tables of e^{-x} and thus find the value of the integral

$$\int_0^1 e^{-x^2}\,dx.$$

Compare your result with the value given in tables of the probability integral.

6. A unit mass is acted upon by a force whose value at successive time intervals is shown by the table

t	0	1	2	3	4	5	6
F	2.4	1.8	1.3	0.9	0.6	0.4	0.3.

Using the equation $mv = mv_o + \int_o^t F\,dt$, construct the curve for v as a function of t, assuming $v_o = 2$.

7. Using the equation $s = s_o + \int_o^t v\,dt$, find the curve for s in terms of t in problem 6, taking $s_o = 1$.

8. From 6 and 7 construct a curve for F as a function of s, then using $U = \int_{s_o}^{s} F\,ds$ find U at $t = 6$ and check by the law

$$U = \frac{1}{2}m(v^2 - v_o^2).$$

9. Check the value of U again from 6 by the equation

$$U = \int_o^t F\,v\,dt.$$

10. A beam 12 feet long is loaded symmetrically, and supported at the two ends. The load in tons per foot run at x feet from either end is given by the table

x.....	0	1	2	3	4	5	6
w.....	0.9	1.5	2.0	2.4	2.7	2.9	3.0

Draw the shearing force and bending moment curves for six feet of the beam.

11. The amount of earth to be removed from a railroad cut is to be computed from the following data, giving cross-sectional area at distances x along the road.

x (yds.).....	0	10	20	30	40	50	60
A (sq. yds.)...	0	11	43	57	62	35	0

Find the number of cubic yards to be moved and the average distance of this earth from the point where x = o.

12. The area in the figure is divided and dimensioned in proper manner for application of the mid-ordinate rule. Vertical strips are of width 0.5, and horizontal strips of width 0.4. Find the area, the center of gravity, and the moments of inertia with respect to the edges of the outer rectangle.

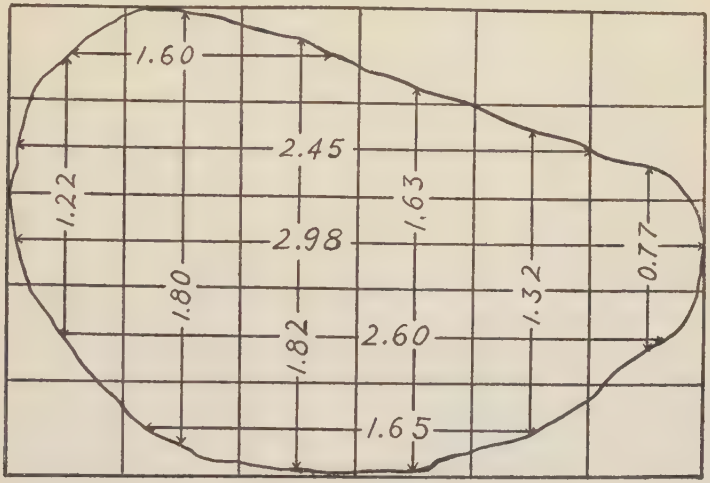


FIG. 82

BOOKS FOR SUPPLEMENTARY READING

Advanced Calculus, W. F. Osgood (Macmillan, 1925), Chapter III.
Mathematical Analysis, Goursat–Hedrick, Vol. I (Ginn, 1904), Chapter VI.
Integral Calculus, W. E. Byerly (Ginn, 1888).
Graphical and Mechanical Computation, Joseph Lipka (Wiley, 1918), Chapter IX.
The Calculus of Observations, Whitaker and Robinson (Blackie and Son, 1924), Chapter VII.

CHAPTER X

METHODS OF EVALUATING FUNCTIONS

Art. 46. Indeterminate Forms

In Chapter II we have given definitions of limits and of certain kinds of infinite variation, under various modes of behavior of the independent variable. We were concerned, also, with the behavior of the combinations $u + v$, uv and u/v when the functions u and v are assumed to approach stated limits or to become infinite in specified manner. In some cases the assumptions as to the behavior of u and v could not yield definite conclusions as to the behavior of the combination of u and v. In such cases we used the expression **indeterminate form,** and stated that further information concerning the functions u and v must be applied to determine the behavior in question. It is the purpose of the present article to show how to apply additional information to evaluate indeterminate forms, not only of the types arising in Chapter II, but also of other types that may be met in practice.

In the following paragraphs we find application for the general theorem stated at the end of Art. 7. It will be convenient at this point to state as a separate theorem that part of the general theorem which we now wish to use.

Intermediate Value Theorem. If $f(x)$ is continuous on the interval from $x = a$ to $x = b$, and if k is between $f(a)$ and $f(b)$, then there is at least one value of x, say c, between a and b, such that $f(c) = k$.

The following proposition follows obviously from the theorem.

Corollary. If $f(x)$ is continuous from a to b, and if $f(x)$ has values both larger and smaller than k between a and b, then for some value c between a and b, $f(c) = k$.

It was this proposition that enabled us to establish Rolle's theorem in Art. 37.

1. **The law of the mean.** The mean value y_m of a function $y = \phi(x)$ between a and b is given by the equation

$$y_m = \frac{\int_a^b \phi(x)\,dx}{\int_a^b dx} = \frac{f(b) - f(a)}{b - a},$$

where $f(x)$ is an integral of $\phi(x)$. The numerator here is, of course, the area under the graph (Fig. 83) of the function between a and b, and the denominator is the distance from a to b. The mean value

of y, the ordinate of the graph, is the altitude of the rectangle of length b-a whose area is equal to the given area.

If the function $\phi(x)$ is continuous on the interval from a to b, then, since the mean value is between the greatest and least values, we see by the corollary above that at some point c between a and b the function is equal to its mean value,

$$\phi(c) = y_m = \frac{f(b) - f(a)}{b - a}.$$

The function $\phi(x)$ being the derivative of f(x), the discussion leads to the following proposition:

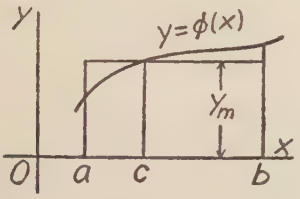


Fig. 83

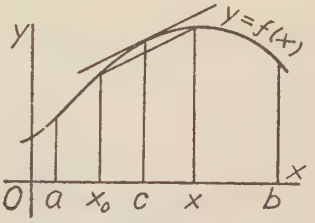


Fig. 84

Mean value theorem. If f(x) and its derivative f′(x) are continuous on the interval form x = a to x = b, then, there is a value c between a and b such that

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

If x_0 and x are any two values between a and b, then the function and its derivative are continuous also from x_0 to x, and these may be substituted for a and b in the conclusion of the proposition. We have then the corollary,

Corollary. If f(x) and f′(x) are continuous from a to b, and if x_0 is between a and b, then for every x between a and b the equation

$$f(x) = f(x) + (x - x_0) f'(c),$$

is true for some c between x_0 and x.

The significance of this proposition is shown geometrically in the figure (Fig. 84). The corollary is equivalent to the statement that the slope of the secant,

$$\frac{f(x) - f(x_0)}{x - x_0},$$

is equal to the slope of the curve at some point c between x_0 and x.

The chief importance of these propositions lies in their usefulness in developing other theorems that follow, but it is well here to note a simple application of the corollary just stated. Suppose the values $f(x_o)$ and $f'(x_o)$ known, but other values of the function unknown, and let it be required to find values of $f(x)$ for x near x_o. The value c is not known, but the value of $f'(c)$, the slope of the secant, will not differ much from $f'(x_o)$, if x is near x_o.

An approximate value of $f(x)$ may be obtained, therefore, by the formula

$$f(x) = f(x_o) + (x - x_o)\, f'(x_o).$$

For example, let $f(x) = \log x$, then $f'(x) = \frac{1}{x}$. We know $\log e = 1$, therefore,

$$\log x = 1 + (x - e)\frac{1}{e} = \frac{x}{e}.$$

This formula will give good results if x is near e, but how near it must be for a given degree of accuracy we are not yet in position to judge.

It should be noticed that the approximate equation used above, where $f'(x_o)$ is used in place of the unknown $f'(c)$, is identical with the equation of the tangent line to the curve $y = f(x)$ at x_o,

$$y - f(x_o) = f'(x_o)\,(x - x_o),$$

as given in Art. 36. Of course the tangent line is very near the curve at points in the neighborhood of the point of tangency.

An important use of this equation is to estimate the error in the value of a function due to an error in the value of the variable. For example, the law of sines for a plane triangle gives

$$a \sin B = b \sin A.$$

Suppose A, B and b are measured to find a, and we wish to know what error will result in a from an error of given magnitude in A, when the measured values are A_o, B_o and b_o. Regarding a as a function of A, we have the approximate equation

$$a - a_o = \frac{da_o}{dA_o}(A - A_o),$$

or, denoting $a - a_0$ and $A - A_0$ by Δa and ΔA and computing the derivative at the given values of A, B and b,

$$\Delta a = \frac{b_0 \cos A_0}{\sin B_0} \Delta A.$$

The equation reveals the facts that the error Δa resulting from a given error ΔA depends on the values b_0, A_0, and B_0, and that it will be extremely large if B_0 is small. The approximate Δa given here is in reality the differential da corresponding to the increment $\Delta A = dA$ in the independent variable.

2. **The Double Law of the Mean.** We wish now to show that if $f(x)$ and $\phi(x)$ are continuous on the interval from $x = a$ to $x = b$, and have continuous derivatives on this interval, then there is a value c between a and b such that

$$\frac{f(b) - f(a)}{\phi(b) - \phi(a)} = \frac{f'(c)}{\phi'(c)}.$$

We may show this by use of the function

$$[f(b) - f(a)]\,[\phi(x) - \phi(a)] - [\phi(b) - \phi(a)]\,[f(x) - f(a)].$$

This function is continuous on the interval from a to b, and is zero at $x = a$ and at $x = b$, therefore, by the mean value theorem, its derivative will be zero at some value, $x = c$, between a and b. Taking the derivative and putting c for x we have

$$[f(b) - f(a)]\,\phi'(c) - [\phi(b) - \phi(a)]\,f'(c) = 0.$$

This equation reduces at once to the relation desired.

We may state a theorem analogous to the corollary in 1 above as follows:

Theorem: **If $f(x)$ and $\phi(x)$ and their first derivatives are continuous on the interval from $x = a$ to $x = b$, and if x_0 is on the interval from a to b, then for every x of the interval there is a c between x_0 and x so that**

$$\frac{f(x) - f(x_0)}{\phi(x) - \phi(x_0)} = \frac{f'(c)}{\phi'(c)}.$$

3. **The form $\frac{0}{0}$.** Suppose that the functions $f(x)$ and $\phi(x)$ are continuous and have continuous derivatives on an interval including x_0,

and suppose that both functions are zero at x_o. The function which is the ratio of these functions

$$\frac{f(x)}{\phi(x)}$$

takes the form $\frac{0}{0}$ at $x = x_o$, and we wish to find how to assign it a value at this point. By the double law of the mean we have

$$\frac{f(x) - f(x_o)}{\phi(x) - \phi(x_o)} = \frac{f'(c)}{\phi'(c)},$$

where c is between x and x_o. Now, $f(x_o)$ and $\phi(x_o)$ are both zero, hence

$$\frac{f(x)}{\phi(x)} = \frac{f'(c)}{\phi'(c)},$$

and as x approaches x_o the variable c does also, and, the two ratios being always equal, they must approach the same limit or become infinite in the same sense. We see, then, that

$$\lim_{x \to x_o} \frac{f(x)}{\phi(x)} = \lim_{c \to x_o} \frac{f'(c)}{\phi'(c)} = \lim_{x \to x_o} \frac{f'(x)}{\phi'(x)}.$$

For this last equality we merely have a different letter, x for c, which can not affect the limit.

As an example we may write

$$\lim_{x \to o} \frac{1 - \cos x}{\sin x} = \lim_{x \to o} \frac{\sin x}{\cos x} = 0.$$

We may, in this case, check the result by a transformation,

$$1 - \cos x = 2 \sin^2 \frac{x}{2}, \quad \sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2},$$

from which we have

$$\lim_{x \to o} \frac{1 - \cos x}{\sin x} = \lim_{x \to o} \frac{\sin \frac{x}{2}}{\cos \frac{x}{2}} = 0.$$

In applying this rule, if the derivatives of numerator and denominator both vanish, then the new function is of the form $\frac{0}{0}$, and the rule may be applied again, taking the ratio of the second derivatives. If

these are both zero, apply the rule again, and so on until either numerator or denominator or both fail to be zero at the given point. At each stage desired algebraic or trigonometric transformation may be used to simplify the function whose limit is sought.

4. **The form** $\frac{\infty}{\infty}$. Suppose now that $f(x)$ and $\phi(x)$ are continuous and have continuous derivatives at one side of x_1, but that the graphs of both functions have vertical asymptotes at x_1. Now the fraction $f(x)/\phi(x)$ takes the form $\frac{\infty}{\infty}$ at x_1, i.e., both numerator and denominator

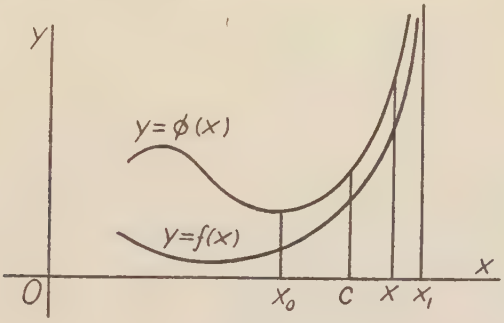


FIG. 85

increase indefinitely as x approaches x_1. We may arrive at a definite value to assign the ratio at this point, or show that it becomes infinite, by an application of the double law of the mean. Taking x_0 and x near to x_1 (Fig. 85) where the functions and their derivatives are continuous, and x between x_0 and x_1, we have

$$\frac{f(x) - f(x_0)}{\phi(x) - \phi(x_0)} = \frac{1 - \dfrac{f(x_0)}{f(x)}}{1 - \dfrac{\phi(x_0)}{\phi(x)}} \cdot \frac{f(x)}{\phi(x)} = \frac{f'(c)}{\phi'(c)},$$

where c is between x_0 and x. Now consider this relation as x and x_0 are regarded as approaching x_1. The three ratios

$$\frac{1 - \dfrac{f(x_0)}{f(x)}}{1 - \dfrac{\phi(x_0)}{\phi(x)}}, \quad \frac{f(x)}{\phi(x)}, \quad \frac{f'(c)}{\phi'(c)},$$

may receive attention individually. Let us suppose that the third of these, which is a continuous function of c, approaches a limit or be-

comes infinite in a definite manner as c approaches x_1, then the product of the other two ratios must do likewise. For any value of x_0 it is possible to choose x so that the fractions

$$\frac{f(x_0)}{f(x)}, \frac{\phi(x_0)}{\phi(x)}$$

are as small as we please, so that while x_0 approaches x_1, x may approach it enough faster to make these fractions approach zero, and therefore the complex fraction, first of the three ratios, may be made to approach unity. The result of this is that the second and third ratios must approach the same limit, or become infinite in the same sense. We have then, under the assumptions made,

$$\lim_{x \to x_1} \frac{f(x)}{\phi(x)} = \lim_{c \to x_1} \frac{f'(c)}{\phi'(c)} = \lim_{x \to x_1} \frac{f'(x)}{\phi'(x)}.$$

This conclusion is identical in form with that for the form $\frac{0}{0}$, and is applied in the same way, the rule being capable of repeated application if the indeterminate character persists.

The fraction

$$\frac{\tan x}{\log\left(\frac{\pi}{2} - x\right)}$$

takes the form $\frac{\infty}{\infty}$ at $x = \frac{\pi}{2}$. Taking derivatives, we have

$$\frac{\sec^2 x}{\frac{1}{x - \frac{\pi}{2}}} = \frac{\left(x - \frac{\pi}{2}\right)}{\cos^2 x},$$

and this, at $\frac{\pi}{2}$, takes the form $\frac{0}{0}$. Again taking derivatives we obtain

$$\frac{1}{2 \sin x \cos x},$$

and at $x = \frac{\pi}{2}$ this takes the form $\frac{1}{0} = \infty$, i.e., the function becomes infinite at $\frac{\pi}{2}$.

5. **Other indeterminate forms.** Any function that becomes indeterminate at a certain value of x, as x_0, may be treated by its relation to a function of the form $\frac{0}{0}$ or $\frac{\infty}{\infty}$. We indicate briefly here the mode of treatment for the forms that arise.

(a) The form $0 \cdot \infty$. If $f(x_0) = 0$, $\phi(x_0) = \infty$, then

$$f(x_0)\,\phi(x_0) = \frac{f(x_0)}{\frac{1}{\phi(x_0)}} = \frac{0}{0},$$

or

$$f(x_0)\,\phi(x_0) = \frac{\phi(x_0)}{\frac{1}{f(x_0)}} = \frac{\infty}{\infty}.$$

(b) The form $\infty - \infty$. If $f(x_0) = \infty$, $\phi(x_0) = \infty$, then we have

$$f(x_0) - \phi(x_0) = \frac{\frac{1}{\phi(x_0)} - \frac{1}{f(x_0)}}{\frac{1}{f(x_0)\,\phi(x_0)}} = \frac{0}{0}.$$

(c) The forms 0^0, ∞^0, 1^∞. The function $f(x)^{\phi(x)}$ has for logarithm the expression $\log f(x)^{\phi(x)} = \phi(x) \log f(x)$. We see therefore, that a function which assumes one of the three forms listed here has a logarithm which assumes one of the forms previously considered. Having found what value to assign to the logarithm, we find at once the correct value for the original function.

6. **Values at $x = \infty$.** We have considered functions that become indeterminate at $x = x_0$ where x_0 is a definite value of x. Frequently we desire to know what value a function approaches as x increases indefinitely. As x increases the function may become infinite in positive or negative sense, or it may approach a definite value, say k. In the latter case, the graph has as a horizontal asymptote, $y = k$. If the function is indeterminate as x becomes infinite, we may, by the methods just considered, reduce the problem to one involving a fractional form,

$$\frac{f(x)}{\phi(x)},$$

where both numerator and denominator become infinite or approach

zero as x increases indefinitely. Let $x = \frac{1}{u}$, then u approaches zero as x becomes infinite, and

$$\frac{d}{du} f(x) = \frac{d}{du} f\left(\frac{1}{u}\right) = -\frac{1}{u^2} f'\left(\frac{1}{u}\right) = -x^2 f'(x)$$

and similarly

$$\frac{d}{du} \phi(x) = -x^2 \phi'(x).$$

Now

$$\lim_{x \to \infty} \frac{f(x)}{\phi(x)} = \lim_{u \to 0} \frac{f(x)}{\phi(x)} = \lim_{u \to 0} \frac{x^2 f'(x)}{x^2 \phi'(x)} = \lim_{x \to \infty} \frac{f'(x)}{\phi'(x)}.$$

We see, therefore, that limits at $x = \infty$ are subject to the same treatment as limits at any finite value x_0.

EXERCISES XLIII

1. What per cent of error occurs in determining the polar moment of inertia of a section of a solid round shaft about its axis due to an error of 0.1 per cent in determining the diameter of the shaft? $\left(I = \frac{\pi d^4}{32}\right)$.

2. What is the error in determining the sine of an angle from its cosine, if the cosine is determined with an error of .0002, (a) when the cosine is found to be .9927, (b) when the cosine is found to be .3279?

3. The angle of elevation of a mountain peak measured from a point in the plain 5 miles distant from it is $5° 20' \pm 2'$, what is the uncertainty in the computed height of the peak?

4. When a pendulum of length L feet swings through a small angle, θ radians, the time, t seconds, of one swing is

$$t = \pi \sqrt{\frac{L}{g}} \left(1 + \frac{\theta^2}{16}\right).$$

Suppose $\theta = 5°$, L = 36 in. and g = 32.16 ft./sec^2. What is the effect on t of changing θ to 6°? Of changing L to 37 in.? Of changing g to 32.2 ft./sec^2?

5. Show that (a) $\left[\frac{a^x - 1}{b^x - 1}\right]_{x=0} = \log_b a$;

(b) $\left[\frac{x - \tan^{-1} x}{x - \sin^{-1} x}\right]_{x=0} = -2$; (c) $\left[\frac{x(x-1)^n - 2}{x^2 - 2x}\right]_{x=2} = n + \frac{1}{2}$.

6. Evaluate the following:

(a) $\dfrac{\tan nx - n\tan x}{n\sin x - \sin nx}$ at $x = 0$; (b) $\dfrac{xe^x - \log(x+1)}{x^2}$ at $x = 0$;

(c) $\dfrac{x - \sin x}{x^2}$ at $x = 0$; (d) $\dfrac{\tan\theta}{\tan 3\theta}$ at $\theta = \pi$.

7. Assign appropriate values to the following functions at the places indicated.

(a) $(a^x + x)^{\frac{1}{x}}$ at $x = 0$; (b) $(\tan\theta)^{\cos\theta}$ at $\theta = \dfrac{\pi}{2}$;

(c) $\dfrac{x^2}{e^x}$ at $x = \infty$; (d) $\dfrac{\cot x}{\cot 3x}$ at $x = 0$;

(e) $\dfrac{\log x}{x}$ at $x = \infty$; (f) $\dfrac{e^x}{\log x}$ at $x = \infty$.

8. Determine the following limits:

(a) $\lim\limits_{x\to\infty} \dfrac{2 - 3x + 5x^5}{7x + x^3 + 7x^5}$; (b) $\lim\limits_{x\to 0} (\cot x)^x$;

(c) $\lim\limits_{x\to\infty} \left(\dfrac{a}{x} + 1\right)^x$; (d) $\lim\limits_{x\to\frac{\pi}{2}} (x\tan x - \dfrac{\pi}{2}\sec x)$.

9. Show that the integral

$$\int_0^\infty xe^{-x}\,dx$$

exists and find its value.

10. Show that $x^n e^{-kx}$ approaches zero as x becomes infinite, for any positive integral value of n, and any positive value of k.

Art. 47. Taylor's Theorem, With Remainder

1. **Statement and proof of theorem.** In Art. 46 we considered the Mean Value Theorem, and we saw how it may be used to approximate the value of a function in the neighborhood of a point where the values of the function and its derivative are known. We are now to derive a theorem of the same character, but much more useful. The theorem may be stated as follows:

Taylor's Theorem. If $f(x)$ and its first n derivatives are continuous

on the interval from $x = a$ to $x = b$, then if x_o is on the interval the equation

$$f(x) = f(x_o) + (x - x_o) f'(x_o) + \frac{(x - x_o)^2}{2!} f''(x_o) + \ldots .$$

$$+ \frac{(x - x_o)^{n-1}}{(n-1)!} f^{(n-1)}(x_o) + \frac{(x - x_o)^n}{n!} f^{(n)}(c).$$

is true for any value of x on the interval and for a value of c between x_o and x.

The last term here is different in form from the others, since the derivative is evaluated at the undetermined point c instead of the known point x_o. This last term is known as the **"remainder."** If $n = 1$, this theorem is identical with the corollary to the Mean Value Theorem.

We may indicate the proof of Taylor's Theorem by treating the case when $n = 3$, the discussion being obviously applicable for any value of n. Let $f(x)$, $f'(x)$, $f''(x)$ and $f'''(x)$ be continuous on the interval from $x = a$ to $x = b$, and let $\overline{M}_3$ and $\underline{M}_3$ be respectively the greatest and least values of $f'''(x)$ on the interval. We have

$$\underline{M}_3 \leqq f'''(x) \leqq \overline{M}_3$$

for every x on the interval, therefore

$$\int_{x_o}^{x} \underline{M}_3 dx \leqq \int_{x_o}^{x} f'''(x)dx \leqq \int_{x_o}^{x} \overline{M}_3 dx,$$

if x is greater than x_o, both x and x_o being on the interval. If x is less than x_o, we should merely reverse the order of the limits of integration, or reverse the inequality signs. We proceed supposing x greater than x_o. Integrating,

$$\underline{M}_3(x - x_o) \leqq f''(x) - f''(x_o) \leqq \overline{M}_3(x - x_o).$$

These functions may be integrated again between the limits x_o and x, and the integrals will again fulfill the inequalities, since if one curve lies below another, the area under it is less than that under the other. Thus, by integrating a third time,

$$\underline{M}_3 \frac{(x - x_o)^3}{3!} \leqq f(x) - f(x_o) - (x - x_o) f'(x_o) - \frac{(x - x_o)^2}{2} f''(x_c)$$

$$\leqq \overline{M}_3 \frac{(x - x_o)^3}{3!}.$$

Now we clearly have

$$f(x) - f(x_o) - (x - x_o)\, f'(x_o) - \frac{(x - x_o)^2}{2} f''(x_o) = \frac{(x - x_o)^3}{3!} M_3,$$

where M_3 is between $\underline{M}_3$ and $\overline{M}_3$. If we had supposed x less than x_o, the inequalities would have reversed each time integrals were taken, but this final equation would have been reached in the same form. Now, for any choice of x and x_o, all that is required of a and b is that x and x_o, shall be on the interval from a to b, and this will be fulfilled if this interval coincides with the interval from x_o to x, therefore this final equation is true if $\underline{M}_3$ and $\overline{M}_3$ are the least and greatest values of $f'''(x)$ on the interval from x_o to x. By the Intermediate Value Theorem, then, there is a c between x_o and x such that $M_3 = f'''(c)$, and the conclusion may be written

$$f(x) = f(x_o) + (x - x_o)\, f'(x_o) + \frac{(x - x_o)^2}{2!} f''(x_o) + \frac{(x - x_o)^3}{3!} f'''(c),$$

where c is between x_o and x. Note that M_3 and c depend on x_o and x.

Another form very frequently used for Taylor's Theorem is reached by putting $x = x_o + h$, obtaining

$$f(x_o + h) = f(x_o) + hf'(x_o) + \frac{h^2}{2!} f''(x_o) + \frac{h^3}{3!} f'''(x_o + \theta h)$$

where θ is between zero and unity. This equation may be extended to any number of terms, the last term always involving the value of the highest derivative at $x_o + \theta h$, and being called the remainder.

2. Use of the theorem. In using the theorem just proved, there are three things at our disposal, (a) the value x_o at which the function and its successive derivatives are to be evaluated; (b) the place x at which the value of the function is considered, or, in other words, the magnitude of $h = (x - x_o)$ whose successive powers occur; (c) the number of terms to be used. The value of c is not known in practice, so that the remainder term is not used in numerical work, except that $f^{(n)}(c)$ may be known to lie between certain values so that in neglecting the remainder we may be sure that we are securing results that are sufficiently accurate.

If the remainder term is neglected, the effect of the formula is to substitute a polynomial of degree $n-1$ for the function f(x), and this polynomial will be a good representation or approximation for the

function if the neglected remainder is small. The theorem is useful, then, only when we can exercise some control over this remainder, and this is done by control over the three things at our disposal mentioned above.

To illustrate take $f(x) = \log x$. Now the choice of x_0 will depend on the region in which it is desired to treat the function, and must be at a point where the function and its derivatives can be evaluated. We cannot take x_0 negative, for $\log x$ is imaginary for negative values. At $x = 0$, $\log x$ and all its derivatives are discontinuous, so that x_0 must be some positive number. Let us take $x_0 = 1$. Computing derivatives and evaluating at $x = 1$,

$$f(x) = \log x, \qquad f(1) = 0;$$

$$f'(x) = \frac{1}{x}, \qquad f'(1) = 1;$$

$$f''(x) = -\frac{1}{x^2}, \qquad f''(1) = -1;$$

$$f'''(x) = \frac{2}{x^3}, \qquad f'''(1) = 2;$$

. . . . . .

$$f^{(n)}(x) = (-1)^{n-1}\frac{(n-1)!}{x^n}; \quad f^{(n)}(1) = (-1)^{n-1}(n-1)!$$

The formula gives

$$\log x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} + \cdots + (-1)^{n-1}\frac{(x-1)^n}{nc^n},$$

where c is between 1 and x. The last term is the remainder, and we have no means of finding its value, but for a given value of x we can state limits which it cannot exceed, having selected a value of n. Take $n = 4$, then, if $x = 2$, the greatest value of the remainder corresponds to the least value of c, which is certainly not less than 1, hence the remainder is not greater numerically than

$$\frac{(2-1)^4}{4} = \frac{1}{4},$$

This, of course, is too large to neglect, so the formula is not suitable at that value of x. Now take $x = 1.1$, then again, c cannot be less

than 1 and the error in neglecting the remainder cannot be greater numerically than

$$\frac{(1.1-1)^4}{4} = .000025$$

and for values of x nearer 1 the error will be less than this. If $x = .9$, c is less than 1, but not less than .9, and the error will not be greater than

$$\frac{(.9-1)^4}{4(.9)^4} = .000038.$$

In general, the magnitude of the error introduced by neglecting the remainder will decrease as n is taken larger, and will decrease as the difference $x - x_0$ is taken smaller. It depends somewhat upon the choice of x_0 but the nature of this dependence is different for different functions?

EXERCISES XLIV

1. By application of Taylor's theorem, find a polynomial of the second degree in x which is a good approximation for cos x near $x = 0$. Estimate the error in using this approximation when $x = 0.3$ radians. (Take $x_0 = 0$, and take $n = 4$, since, in this case $f'''(x_0) = 0$.)

2. Express $x^3 - 6x^2 + 1$ in powers of $x - 2$. (Take $x_0 = 2$, and $n = 3$, then check by reducing the result to the original form by algebra.)

3. Express sin x in powers of $x - \frac{\pi}{4}$, and show that the remainder after three terms, when $x - \frac{\pi}{4} = .05$, is less than .00005. (Take $x_0 = \frac{\pi}{4}$, and take $n = 3$.)

4. Show by Taylor's theorem that sin x does not differ from x by more than $\frac{x^3}{6}$, and therefore, if $x = 0.1$, the difference is beyond the third decimal place. (Take $x_0 = 0$, $n = 3$.)

5. Find a polynomial approximation for e^x that will be accurate to the third decimal place if $x = \frac{1}{2}$.

6. Find an approximation for $\log_{10}(5 + x)$ in powers of x, accurate to four places if $x = 1$.

7. Show that $1 + x + x^2$ is a good approximation for $\frac{1}{1-x}$, if x is small.

8. Show that $\frac{1}{1+x}$ may be replaced by $1 - x + x^2$, with results accurate to three places if $x = .05$.

9. Show that $1 + \frac{x^2}{2}$ is a proper approximation for $\sqrt{1+x^2}$, if x is small.

10. Show that $8 - 3x + \frac{3}{16}x^2$, is a proper approximation for $(4-x)^{\frac{3}{2}}$, if x is small.

11. Write an algebraic expression for e^x in powers of x that is accurate to six decimal places when $x = 1$. Show that $e = 2.718282-$.

Art. 48. Infinite Series

1. **Taylor's series.** By use of Taylor's theorem we may approximate certain types of functions in the neighborhood of a point x_o by a polynomial, in powers of $x - x_o$. And we have seen that the error depends on the number of terms, i.e., the degree of the polynomial, and on the value of $x - x_o$. Let us denote the error, the remainder after n terms, by $R_n(x)$, and the polynomial of n terms by $P_n(x)$, so that

$$f(x) = P_n(x) + R_n(x).$$

It may be that for some values of x, say for x on the interval from a to b, the remainder $R_n(x)$ approaches zero as n increases indefinitely. If this is true, then the polynomial $P_n(x)$ approaches $f(x)$ for every x on the interval, as n increases indefinitely. The notation employed in such a case is

$$f(x) = f(x_o) + (x - x_o)\,f'(x_o) + \frac{(x-x_o)^2}{2!}\,f''(x_o) + \ldots .$$

the dots indicating that the expression continues indefinitely. The expression on the right is a Taylor's series, and on the interval from a to b, at every point of which we have

$$\lim_{n\to\infty} R_n(x) = 0, \qquad \lim_{n\to\infty} P_n(x) = f(x),$$

the series is an exact representation of the function $f(x)$. A function so expressed is said to be "expanded in powers of $(x - x_o)$" or "developed in a Taylor's series about x_o."

To illustrate, consider $f(x) = e^x$. By Taylor's theorem we have,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{n-1}}{(n-1)!} + \frac{x^n}{n!} e^c,$$

where c is between zero and x. If x is positive, c is less than x, and e^c is less than e^x, while if x is negative, c is also negative and e^c is less than unity. In any case then, the remainder $R_n(x)$ is not greater than

$$\frac{x^n}{n!} e^{|x|}.$$

For a given value of x we have

$$\lim_{n \to \infty} R_n(x) \leqq \lim_{n \to \infty} \frac{x^n}{n!} e^{|x|} = e^{|x|} \lim_{n \to \infty} \frac{x^n}{n!}.$$

This last limit is zero, for

$$\frac{x^n}{n!} = \frac{x}{1} \cdot \frac{x}{2} \cdot \frac{x}{3} \cdots \frac{x}{n},$$

and if m is numerically greater than 2x, all factors after $\frac{x}{m}$ are less than $\frac{1}{2}$, and the product of these is less than $\left(\frac{1}{2}\right)^{n-m}$, which approaches zero as n increases indefinitely. We find then that e^x is perfectly represented by a Taylor's series for all values of x. The polynomial formed of the first few terms of the series is a good approximation for the function, the number of terms required for a certain limit of error depending on the value of x.

Again, by Taylor's theorem,

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots x^{n-1} + \frac{x^n}{(1-c)^{n+1}}.$$

Here, if x lies between $\frac{1}{2}$ and $-\frac{1}{2}$, c does also, and $1 - c$ is greater than $\frac{1}{2}$, and the remainder,

$$R_n(x) = \frac{1}{1-c}\left(\frac{x}{1-c}\right)^n,$$

approaches zero as n increases indefinitely, since $\frac{x}{1-c}$ is less than unity. The Taylor's series therefore represents the function perfectly on the interval from $-\frac{1}{2}$ to $\frac{1}{2}$, but outside of this interval the uncertainty of c leaves us in doubt. But the function becomes infinite at $x = 1$, therefore, x must be less than 1 in order that Taylor's theorem be applicable.

2. **Convergence of series.** In general we may denote a series by

$$S = u_0 + u_1 + u_2 + \ldots u_{n-1} + u_n + \ldots,$$

the notation indicating that,

$$S = \lim_{n \to \infty} S_n,$$

where S_n is the sum of the first n terms of the series. Given a series of terms, it is important to know whether this limit actually exists, i.e., whether S_n approaches a definite value as n increases indefinitely. If this limit exists, the series is said to **converge** or to be **convergent,** otherwise it is said to **diverge,** or be **divergent.** Some series may be shown to be convergent by simple tests, and some may be shown to be divergent by simple tests, but some series require more elaborate tests to determine their character, and some fail to respond to any known test.

We give here a few simple tests to determine the character of a series.

(1) **Positive terms, sum bounded. If the terms of the series**

$$u_0 + u_1 + u_2 + \ldots$$

are all positive, and if S_n, the sum of the first n terms, is less than a certain fixed number, say A, for all values of n, then the series is convergent and

$$\lim_{n \to \infty} S_n = S \leqq A.$$

To establish this proposition, we observe that S_n cannot approach a limit greater than A without itself passing A, then it remains to show that a limit exists. Since all terms are positive, S_n increases with n. Now as n increases some numbers are reached and exceeded by S_n, while others are never reached. All positive numbers belong to one group or the other, and some number, N, will be at the division point

between numbers that are reached and passed and those that are not. This number N is clearly the limit of S_n, and is the value of S.

(b) **Comparison test. Let**

$$u_o + u_1 + \ldots .$$

be a series of positive terms, then if a convergent series of positive terms

$$A = a_o + a_1 + \ldots .$$

can be found such that each term a_n is at least as great as the corresponding term u_n, then the first series converges, and its value S is not greater than A.

If a divergent series

$$b_o + b_1 + b_2 + \ldots .$$

can be found such that each term b_n is not greater than the corresponding term u_n, then the first series is divergent.

For the proof, first as to convergence, we have S_n always less than or equal to the corresponding sum A_n but A_n always less than A, therefore $S_n < A$, then by (a) it is convergent, and its value is $S \leqq A$.

As to divergence, the sum B_n of the first n terms of the divergent series, cannot remain less than any fixed number, for then the series would be convergent by (a). Therefore B_n increases indefinitely, and S_n, which is not less than the corresponding B_n, must also increase indefinitely.

(c) **The test-ratio test. Let**

$$u_o + u_1 + \ldots .$$

be an infinite series, of positive terms, then if

$$\lim_{n \to \infty} \frac{u_{n+1}}{u_n} = t,$$

the series is convergent if $t < 1$, and is divergent if $t > 1$.

The ratio of a term u_{n+1} to the preceding term u_n is called the **test ratio.** If its limit, t, is equal to unity, the test is not effective.

To establish the proposition we use the comparison test, comparing the series

$$u_o + u_1 + u_2 + \ldots .$$

with the series of positive terms

$$a + ar + ar^2 + \ldots .$$

By algebra we have

$$\frac{a(1 - r^n)}{1 - r} = a(1 + r + r^2 + \ldots . r^{n-1}).$$

Now the first member here may be written

$$\frac{a}{1 - r} - \frac{ar^n}{1 - r}.$$

If $r < 1$, the second term here approaches zero as n increases indefinitely, while if $r \geqq 1$, the limit does not exist. We see, then, that

$$\frac{a}{1 - r} = a + ar + ar^2 + \ldots ,$$

is a convergent series if r is numerically less than unity, otherwise it is divergent.

Now suppose $t < r < 1$, then, since the test ratio approaches t, there is some number, m, such that if $n > m$ then

$$\frac{u_{n+1}}{u_n} < r.$$

Take $ar^m = u_m$, then, for $n > m$, we have $ar^n > u_n$, therefore, since the first m terms cannot affect the convergence, the series converges by the comparison test.

Now suppose $1 \leqq r < t$, then, reasoning as before, for some m we have the test ratio greater than r for $n > m$, and the given series may be shown to diverge by the comparison test.

(d) **Absolute convergence. A series whose terms are of different signs is convergent if the series formed by making all terms positive is convergent.**

A series which responds to this test is said to be **absolutely convergent.** Any convergent series of positive terms is absolutely convergent.

To prove the proposition stated above, let S_n be the sum of the first n terms all considered positive, and let A_n be the sum of those among the first n terms that were positive in the original series, and B_n the

sum of those that were originally negative. If T_n is the algebraic sum of the first n terms of the original series we have

$$T_n = A_n - B_n, S_n = A_n + B_n.$$

As n increases indefinitely, S_n approaches a limit S, therefore A_n and B_n must approach limits A and B, and consequently T_n approaches the limit $T = A - B$.

(e) **Alternating series. If the terms of a series**

$$u_o - u_1 + u_2 - \ldots .$$

are alternately positive and negative, and if

$$|u_n| \geqq |u_{n+1}|$$

for all values of n, and if

$$\lim_{n \to \infty} u_n = 0$$

then the series converges.

Here the sum S_n of the first n terms alternately increases and decreases as n increases, the changes growing less and finally approaching zero. Let $\bar{S}_n$ be the greater of S_n and S_{n-1}, and let $\underline{S}_n$ be the smaller of these, then $\bar{S}_n = \underline{S}_n + u_{n-1}$. Now $\underline{S}_n$ increases with n, and never exceeds u_o, hence, reasoning as for test (a) above, it must approach a limit $\underline{S}$. Similarly, $\bar{S}_n$ decreases with n, and is never less than $u_o - u_1$, therefore it must approach a limit $\bar{S}$. Since u_{n-1} approaches zero, $\bar{S} = \underline{S}$, and since $\underline{S}_n$ is always either $\bar{S}_n$ or $\underline{S}_n$, it must approach a limit $S = \bar{S} = \underline{S}$.

We notice that if the terms of a series that converges by this test are all made positive, the test ratio of the new series is less than 1, but may approach 1 as a limit. The series is therefore not necessarily absolutely convergent. A series that converges but not absolutely is called **conditionally** convergent. An important distinction between the two kinds of convergent series is that a change in the order of the terms of an absolutely convergent series cannot affect its convergence, while rearrangement of a conditionally convergent series may produce a divergent series.

(f) **Test by integration.** The nth term of a series is, in general, a function of n, say $u_n = f(n)$, where n is regarded as taking only positive integral values. Now **if we regard n as a continuous variable,**

and if the derivative f'(n) is negative for all values of n and if the integral

$$\int_0^\infty f(n)\, dn$$

exists, then the series is absolutely convergent and the sum S satisfies the condition

$$S - u_0 < \int_0^\infty f(n)\, dn < S.$$

If the integral fails to exist (is infinite) then the series is divergent.

We may see the truth of this by reference to a graph (Fig. 86). Let the equation of the curve be $u = f(n)$, then the area between the

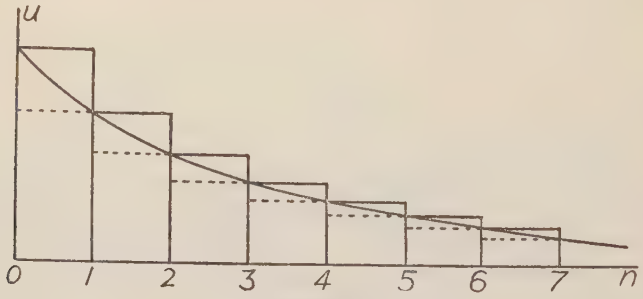


Fig. 86

curve and the n-axis is given by the integral, while the sum of the areas of the full line rectangles is S and the area under the dotted lines is $S - u_0$. Since the slope of the curve is known to be negative the relations must be as shown.

(g) **Use of tests.** The tests for convergence and divergence discussed here are far from exhaustive, but include those most likely to be useful and convenient in practice. No general rule can be stated as to which test to try first in practice, but the character of the given series will usually suggest a choice of tests. Convenience of presentation determined the order of the tests (a) to (f), and this order is quite different from the order of practical utility. In fact (a) would rarely be used directly in practice, and (d) involves the use of another test. Often (c) is tried first, and if it fails (b) or (f) may be tried. The test (e) is only for conditional convergence, and must be followed by another test to determine the question of absolute convergence.

In testing a series by any method it is well to remember that the convergence or divergence of a series is not changed by the omission of any finite set of terms, nor by multiplying all terms of the series by a common factor. This is practically important in applying the comparison test, and also in dealing with alternating series.

As an example, consider the series

$$x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \ldots$$

The test ratio is $x^2 \dfrac{2n+1}{2n+3}$, and the limit of this is x^2, therefore the series is absolutely convergent for $|x| < 1$ and divergent for $|x| > 1$, but the test is not effective when $x = \pm 1$. If $x = \pm 1$ the signs alternate, and the series is convergent by test (e). The convergence is conditional, however, for if we make all signs positive and apply (f) we have

$$\int_0^\infty \frac{dn}{2n+1} = [\log \sqrt{2n+1}]_0^\infty = \infty,$$

showing that the series is not absolutely convergent at $x = 1$ or $x = -1$.

EXERCISES XLV

1. Find the Taylor's series for sin x in powers of x. Show by the test ratio that the series converges for all values of x. By Taylor's theorem with remainder, estimate the number of terms necessary to give $\sin \frac{\pi}{7}$ accurate to five places. Estimate the number of terms necessary to give $\sin \frac{5\pi}{7}$ accurate to five places.

2. Show that the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$ is divergent. (Compare with the series

$$\frac{1}{2} + \frac{1}{2} + \left[\frac{1}{4} + \frac{1}{4}\right] + \left[\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right] + \ldots).$$

3. Show that the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

is conditionally convergent.

4. Show by comparison with the harmonic series (example 2) that the series

$$1 + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \ldots .$$

is divergent if $p \leqq 1$. Also test by integration.

5. Show that the series in example 4 is convergent if $p > 1$. (Make the comparison

$$1 + \left[\frac{1}{2^p} + \frac{1}{3^p}\right] + \left[\frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p}\right] + \ldots .$$

$$1 + \left[\frac{1}{2^p} + \frac{1}{2^p}\right] + \left[\frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p}\right] + \ldots .$$

and observe that this last series is equal to

$$1 + r + r^2 + r^3 + \ldots .$$

where $r = \frac{1}{2^{p-1}}$. Also test directly by integration.)

6. Obtain the Taylor's series for $(a + x)^p$, where p is any real number, positive or negative, and show, by the test ratio, that the series converges if $|x| < |a|$, but diverges if $|x| > |a|$. Show that if p is a positive integer the series terminates and is identical with the form obtained by the Binomial Theorem (Art. 19). (In this special case, p being any real number, Taylor's series gives what is known as the General Binomial Theorem.)

7. Obtain by use of the binomial theorem, the Taylor's series for $(1 + x^2)^{\frac{2}{3}}$, and show that it is convergent if $|x| < 1$. (Example 6.)

8. Show that the series

$$1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots .$$

converges absolutely if $x < 1$, (test ratio), that it converges conditionally if $x = -1$ (example 13), and that it diverges if $x = 1$ or $|x| > 1$. (Compare with harmonic series.)

9. Test the following series for convergence:

(a) $1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots + \frac{1}{2n+1} + \ldots ;$

(b) $\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \ldots + \frac{1}{2^n + 1} + \ldots ;$

(c) $\frac{1}{10} - \frac{1}{20} + \frac{1}{30} - \frac{1}{40} + \ldots ;$

10. For what values of x are the following series absolutely convergent?

(a) $x - \frac{x^3}{\sqrt{3}} + \frac{x^5}{\sqrt{5}} - \ldots ;$

(b) $1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \ldots ;$

(c) $\frac{1}{1+x} + \frac{1}{1+x^2} + \frac{1}{1+x^3} + \ldots ;$

(d) $1 + \frac{x}{3} + \frac{x^2}{3^2} + \frac{x^3}{3^3} + \ldots ;$

(e) $1 - x + 2x^2 - 3x^3 + \ldots$

Art. 49. Operations on Series

We have seen that if a series whose terms involve a variable, x, converges for every value of x on a certain interval, say from a to b, then the series is a perfect representation of a function of x on the interval. We are now to consider how certain operations on the functions represented by series may be accomplished by operations on the series themselves. First, however, we must have more complete understanding of the character of series and the functions represented by them.

Let,

$$u_0(x) + u_1(x) + u_2(x) + \ldots$$

be a series of functions and suppose that

$$a_0 + a_1 + a_2 + \ldots$$

is a series of constant positive terms such that $|u_n(x)| \leqq a_n$ for every value of n, and for every x on the interval from $x = a$ to $x = b$. Under these conditions the series of constants will be said to **dominate** the series of functions on the interval from a to b. An important proposition is the following:

If a series of functions is dominated by a convergent series of constant terms on an interval from a to b, then the series of functions is absolutely convergent for every x on the interval.

This is evident from the discussion of absolute convergence given above.

Another important proposition is:

If every term of the series

$$u_0(x) + u_1(x) + u_2(x) + \ldots .$$

is a continuous function on the interval from a to b, and if the series is dominated by a convergent series of constants on this interval, then the series represents a continuous function U(x) on the interval.

To see the truth of this let $U_n(x)$ denote the sum of the first n terms. Since the sum of two or more continuous functions is continuous, $U_n(x)$ is continuous on the given interval. That $U_n(x)$ approaches a definite limit for each x on the interval is a result of the previous theorem.

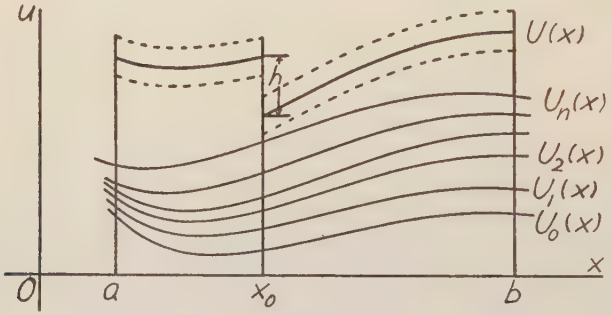


FIG. 87

Take a function U(x) so that for any x on the interval $|U(x)|$ is the limit $U_n(x)$, then it remains to show that the function $U_{(x)}$ so defined is continuous.

Let the dominating series be

$$A = a_0 + a_1 + a_2 + \ldots ;$$

and let A_n be the sum of the first n terms. Since $|U(x)| \leqq A$ for every x on the interval, the function cannot have a vertical asymptote. Now suppose that U(x) were to have a finite discontinuity at some place, x_0 (Fig. 87), and suppose that the magnitude of the break is h, as shown. Now for every x on the interval we have

$$|U(x) - U_n(x)| \leqq A - A_n,$$

since the remainder after n terms of a series is itself a series, and the conditions for absolute convergence apply. But as n increases, $A - A_n$

becomes smaller than $\frac{h}{2}$, so that $U_n(x)$ must be close to $U(x)$, say within the dotted lines in the figure. Since this is impossible for a continuous $U_n(x)$, we see that $U(x)$ cannot be discontinuous.

We now consider conditions under which certain operations on series are valid.

1. **Addition and subtraction. If the two series**

$$U(x) = u_0(x) + u_1(x) + \ldots + U_n(x) + \ldots,$$

$$V(x) = v_0(x) + v_1(x) + \ldots + v_n(x) + \ldots,$$

are absolutely convergent for a certain value of x, then for that value of x we have

$$U(x) \pm V(x) = [u_0(x) \pm v_0(x)] + [u_1(x) \pm v_1(x)] + \ldots + [u_n(x) \pm v_n(x)] + \ldots$$

and this last series is absolutely convergent.

On account of the nature of absolute convergence, the proof will be complete if we can establish the conclusions for addition assuming that each term of both given series is positive for the value of x in question. Let $U_n(x)$ and $V_n(x)$ denote the sum of the first n terms of the respective series and let

$$U(x) = U_n(x) + R_n(x), \quad V(x) = V_n(x) + T_n(x),$$

then

$$U(x) + V(x) = U_n(x) + V_n(x) + R_n(x) + T_n(x).$$

Since $R_n(x)$ and $T_n(x)$ approach zero as n increases indefinitely, the sum $U_n(x) + V_n(x)$ approaches the sum $U(x) + V(x)$. Further, $U_n(x) + V_n(x)$ is the sum of n terms of the new series, and this series therefore converges and represents the sum $U(x) + V(x)$.

2. **Multiplication. If, for a certain value of x, the two series**

$$U(x) = u_0(x) + u_1(x) + \ldots + u_n(x) + \ldots,$$

$$V(x) = v_0(x) + v_1(x) + \ldots + v_n(x) + \ldots,$$

are absolutely convergent, then for that value of x we have

$$U(x) \cdot V(x) = u_0(x) \cdot v_0(x) + u_1(x) \cdot v_0(x) + u_0(x) \cdot v_1(x) + \ldots + u_{n-1}(x) \cdot v_n(x) + u_n(x) \cdot v_n(x) + \ldots,$$

and this new series is absolutely convergent.

Here again, the proof will be complete if we can establish the conclusion assuming that all terms are positive for the value of x considered. Using the notations as before, we have

$$U(x) \cdot V(x) = U_n(x) \cdot V_n(x) + U_n(x) \cdot T_n(x) + V_n(x) \cdot R_n(x) + R_n(x) \cdot T_n(x).$$

Now as n increases indefinitely $R_n(x)$ and $T_n(x)$ approach zero, and $U_n(x) \cdot V_n(x)$ approaches the product $U(x) \cdot V(x)$. The terms of the new series are of the form $u_i(x) \cdot v_j(x)$, and we may denote by S_k the sum of all terms such that $i + j \leqq k$. First, if $k = 2n$, then S_k contains all terms in the product $U_n(x) \cdot V_n(x)$ and some terms of $U_n(x) \cdot T_n(x)$ and $V_n(x) \cdot R_n(x)$, therefore

$$S_{2n} > U_n(x) \cdot V_n(x).$$

Again, if $k = n$, the product $U_n(x) \cdot V_n(x)$ contains every term in S_k and some terms not in S_k, therefore,

$$U_n(x) \cdot V_n(x) > S_n.$$

Now it is obvious that as n increases S_n may be made to reach any magnitude that S_{2n} reaches, and if one of these approaches a limit the other does also. But S_n being less than $U_n(x) \cdot V_n(x)$ must approach a limit less than or equal to $U(x) \cdot V(x)$, and S_{2n} being greater than $U_n(x) \cdot U_n(x)$, cannot have a limit less than $U(x) \cdot V(x)$ therefore the common limit is clearly $U(x) \cdot V(x)$.

3. **Integration. If the series of functions**

$$U(x) = u_0(x) + u_1(x) + \ldots + u_n(x) + \ldots$$

is dominated on the interval from $x = a$ to $x = b$ by the convergent series of positive constants,

$$A = a_0 + a_1 + \ldots + a_n + \ldots,$$

and if every term is a continuous function of x on this interval then

$$\int_{x_0}^{x} U(x)dx = \int_{x_0}^{x} u_0(x)dx + \int_{x_0}^{x} u_1(x)dx + \ldots + \int_{x_0}^{x} u_n(x)dx + \ldots,$$

for every x_0 and x on the interval, and the new series is dominated on this interval by the series

$$A(b - a) = a_0(b - a) + a_1(b - a) + \ldots + a_n(b - a) + \ldots,$$

Using the notation employed above we may write

$$\left| U(x) - U_n(x) \right| \leqq A - A_n,$$

therefore, we have

$$\left| \int_{x_0}^{x} U(x)dx - \int_{x_0}^{x} U_n(x)dx \right| \leqq (A - A_n) \cdot \left| x - x_0 \right|.$$

Now the integral

$$\int_{x_0}^{x} U_n(x)dx$$

is the sum of the first n integrals, terms of the new series, and since $A - A_n$ approaches zero as n increases indefinitely, we see that

$$\lim_{n \to \infty} \int_{x_0}^{x} U_n(x)dx = \int_{x_0}^{x} U(x)dx.$$

That the series of integrals is dominated as stated is obvious from a consideration of the individual terms. It is important to notice that this new dominating series is convergent.

4. **Differentiation. If the series of functions**

$$U(x) = u_0(x) + u_1(x) + \ldots + u_n(x) + \ldots$$

is convergent for every x on the interval from x = a to x = b, and if the series of derivatives

$$\frac{d}{dx} u_0(x) + \frac{d}{dx} u_1(x) + \ldots + \frac{d}{dx} u_n(x) + \ldots,$$

is dominated on the interval from a to b by a convergent series of constants

$$A = a_0 + a_1 + \ldots + a_n + \ldots,$$

and if every derivative is continuous on the given interval, then the series of derivatives is a perfect representation on the interval of the derivative,

$$\frac{d}{dx} U(x).$$

To prove this, we merely observe that under the conditions stated the series of derivatives may be integrated according to the previous

proposition, obtaining the original series plus a series of constants, therefore the function it represents can be none other than the derivative of the function U(x).

5. **Power series.** A power series in terms of a variable x is a series of ascending powers of x, each multiplied by a constant coefficient, as

$$a_o + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots$$

Most of the series with which we have to deal are power series.

An important proposition concerning such series is:

If a power series in x converges for $x = x_o$, it converges absolutely for every x such that $|x| < |x_o|$.

To prove this suppose that the series

$$a_o + a_1x_o + a_2x_o^2 + \ldots$$

is convergent. We may take a number M so that, for any n, $M > |a_nx_o^n|$, then if $|x| < |x_o|$,

$$|a_nx^n| = \left|a_n x_o^n \left(\frac{x}{x_o}\right)^n\right| < M\left|\frac{x}{x_o}\right|^n.$$

But the series

$$M + M\frac{x}{x_o} + M\left(\frac{x}{x_o}\right) + \ldots$$

is absolutely convergent when $|x| < |x_o|$, therefore, under the conditions stated the series

$$a_o + a_1x + a_2x^2 + \ldots$$

is absolutely convergent.

Another important proposition is:

If the series

$$a_o + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots$$

converges, for $x = x_o$, then the series formed by differentiating this term by term,

$$a_1 + 2a_2x + \ldots + na_nx^{n-1} + \ldots,$$

converges for every x such that $|x| < |x_o|$.

To prove this, take x_1 positive, so that $|x| < x_1 < |x_o|$, then the series

$$a_o + a_1x_1 + a_2x_1^2 + \ldots + a_nx_1^n + \ldots$$

converges. We may prove the convergence of the derived series above, supposing x positive, by comparison with this last series. We wish to show that

$$na_n x^{n-1} \leqq a_n x_1^n$$

for all values of n beyond a certain number. This is the same as showing that

$$\lim_{n \to \infty} \frac{nx^{n-1}}{x_1^n} < 1.$$

We may show this by taking the limit of the indeterminate form

$$\lim_{z \to \infty} \frac{z}{k^{z-1}} = \lim_{z \to \infty} \frac{1}{k^{z-1} \log k}$$

and this is zero if $k > 1$. Now

$$\frac{nx^{n-1}}{x_1^n} = \frac{1}{x_1} \left(\frac{n}{\frac{x_1}{x}}\right) n - 1$$

and, x_1 being constant, by taking $k = \frac{x_1}{x}$ we see that as n increases this ratio approaches zero.

Since a power series

$$a_o + a_1 x + a_2 x^2 + \ldots .$$

that converges for a value $x = x_o$, is dominated on the interval from $x = -x_1$ to $x = x_1$, where $|x_1| < |x_o|$, by the convergent series of positive constants,

$$|a_o| + |a_1 x_1| + |a_2 x_1^2| + \ldots ,$$

and since every term of a power series is a continuous function on any interval, we see by previous propositions that a power series represents a continuous function on any interval on which it is convergent, and that the operations of addition, subtraction, multiplication, differentiation and integration may be performed on power series term by term, and the results are valid for the functions represented within the interval for convergence.

We note one more proposition:

If a power series in x represents a function on any interval, then it is the Taylor's series for that function.

Let

$$f(x) = a_o + a_1x + a_2x^2 + \ldots .$$

and let the series converge at $x = x_o$, then we have

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + \ldots .$$

$$f''(x) = 2a^2 + 3!\, a_3x + 4.3a_4x^2 + \ldots .$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$f^{(n)}(x) = n!a_n + (n+1)!\, a_{n+1}x + \frac{(n+2)!}{2!}\, a_{n+2}x^2 + \ldots ,$$

for all values of x such that $|x| < |x_o|$. In particular these relations are true when $x = o$, therefore

$$a_o = f(o),\ a_1 = f'(o),\ 2a_2 = f''(o),\ \ldots ,\ n!\, a_n = f^{(n)}(o),$$

and therefore the original series becomes

$$f(x) = f(o) + xf'(o) + \frac{x^2}{2!} f''(o) + \ldots + \frac{x^n}{n!} f^{(n)}(o) + \ldots .$$

This is a Taylor's series in powers of x.

EXERCISES XLVI

1. Write the Taylor's series for sin x in powers of x, then by differentiation find the series for cos x. By the multiplication of series show that $2 \sin x \cos x = \sin 2x$.

2. Write the power series for $\frac{1}{1-x}$ by the binomial theorem, and thus find by integration the series for $\log (1 - x)$.

3. Find the power series for e^x and e^{-x}, then by addition and subtraction find the series for $\cosh x = \frac{1}{2}(e^x + e^{-x})$ and $\sinh x = \frac{1}{2}(e^x - e^{-x})$. By differentiation of series show that each of these last two functions is the derivative of the other.

4. Show that

$$\int_o^x e^{-x^2}\, dx = x - \frac{x^3}{3} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \ldots .$$

5. Derive the series for $\sin^{-1} x$ by integration of the series for

$$(1 - x^2)^{-\frac{1}{2}}$$

6. Expand by the binomial theorem the function $(1 - x^3)^{-\frac{1}{2}}$, and by integration show that

$$\int_0^{\frac{1}{2}} \frac{dx}{\sqrt{1 - x^3}} = .508.$$

7. Show that if a power series is unchanged by differentiation it is the power series for the function $a_o e^x$, where a_o is the first term of the series.

8. Let $y = a_o + a_1 x + a_2 x^2 + a_3 x^3 + \ldots$, and show that if $\frac{d^2y}{dx^2} = -y$ then $y = a_o \cos x + a_1 \sin x$.

BOOKS FOR SUPPLEMENTARY READING

Advanced Calculus, W. F. Osgood (Macmillan, 1925), Chapters VII and X.
Advanced Calculus, E. B. Wilson (Ginn, 1912), Chapters III, XVI and XVII.
Mathematical Analysis, Goursat-Hedrick, Vol. I (Ginn, 1904), Chapters III, VIII and IX.
Engineering Mathematics, C. P. Steinmetz (McGraw-Hill, 1911), Chapter V.

CHAPTER XI

COMPLEX QUANTITIES AND PERIODIC FUNCTIONS

Art. 50. Complex Quantities

1. **Standard forms for complex numbers.** A **complex number** is a quantity of the form $a + bi$, where a and b are real numbers, rational or irrational, and i is $\sqrt{-1}$. The class of complex numbers includes the real numbers and the **pure imaginary** numbers, the latter being real numbers multiplied by the imaginary unit, i. Examples of complex numbers are

$$2 + \sqrt{-9} = 2 + 3\sqrt{-1} = 2 + 3i, \sqrt{-7} = i\sqrt{7}, 4i, \sqrt{2}.$$

Of these four examples the second and third are pure imaginaries and the fourth is a real number.

The student has encountered complex numbers as the roots of algebraic equations, and has had some experience in adding, subtracting, multiplying and dividing such quantities. It is expedient in performing such operations to use the notation i, applying the process just as to any algebraic symbol, finally reducing the result by the relations

$$i = \sqrt{-1},\ i^2 = -1,\ i^3 = -i,\ i^4 = 1,\ i^5 = i. \qquad (1)$$

It is evident that adding, subtracting and multiplying together two numbers of the form $a + bi$ will always yield a result in this same form. If one such number is to be divided by another, we may multiply both dividend and divisor by the conjugate of the divisor, then express the quotient in this same standard form. Thus,

$$\frac{3 - 8i}{2 + 5i} = \frac{(3 - 8i)(2 - 5i)}{(2 + 5i)(2 - 5i)} = \frac{-34 - 31i}{29} = -\frac{34}{29} - \frac{31}{29}i,$$

which last is the required form.

In employing, the standard form $a + bi$ it is well to note that the pair of real numbers, a and b, determine uniquely a complex number. That is, if $a + bi = c + di$, then $a = c$ and $b = d$. For, transposing, we have $a - c = (d - b)i$, and if both sides of this equation were not zero we should have a real number equal to a pure imaginary, which is absurd.

Given any two real numbers a and b, we may take a real positive number ρ and an angle θ so that

$$a = \rho \cos\theta, \quad b = \rho \sin\theta, \quad a + bi = \rho(\cos\theta + i \sin\theta). \tag{2}$$

Here, ρ being always positive, θ is determined by the equations

$$\sin\theta = \frac{b}{\rho}, \quad \cos\theta = \frac{a}{\rho}$$

where

$$\rho = \sqrt{a^2 + b^2}.$$

To arrive at a third important notation for complex numbers we must first assign a meaning to the expression $e^{i\theta}$. We have seen that the exponential function e^x, where x is real, is expressed by a power series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

Now we shall take as the definition of $e^{i\theta}$ the value that this power series takes when x is replaced by $i\theta$, so that, using (1) above,

$$e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} + \cdots \tag{3}$$

Now each term of this series is a perfectly definite number, real or pure imaginary, and separating the reals from the imaginaries, and remembering the power series for $\sin\theta$ and $\cos\theta$, we have

$$e^{i\theta} = \left(1 - \frac{\theta^2}{2!} + \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \cdots\right) = \cos\theta + i\sin\theta.$$

We may now write

$$a + bi = \rho(\cos\theta + i\sin\theta) = \rho e^{i\theta} \tag{4}$$

where ρ and θ are related to a and b as in (2) above. These three equivalent forms are found very useful, the choice of form depending on the processes to be applied. The positive number ρ is called the **modulus** or **absolute value** of the number, while θ is called the **amplitude** or **argument** of the number.

2. **Geometrical representation of complex numbers.** Real numbers may be represented by distances along an axis, or by segments of a

given line. Now a complex number consists of two parts of different kind, a real part, a, and a pure imaginary part, bi. To represent a complex number geometrically, then, it is convenient to represent the real part by a segment of the line OX_1 called the **axis of reals** (Fig. 88), and to represent the imaginary part by a segment of the line OY, perpendicular to OX, called the **axis of imaginaries.** The sum of these two parts, the number a + bi, may be represented by the diagonal, OP, of the rectangle whose sides are of lengths a and b. The length of this line is ρ and it makes the angle θ with OX. Taking OX and OY as axes, the point P has the rectangular coordinates (a, b), and the polar coordinates (ρ, θ).

Any directed line segment in the plane of OX and OY may represent a complex number, and the number may be written in the standard forms of (4) above, where a and b are the projections on the axes of reals and imaginaries respectively, and ρ is the length of the segment

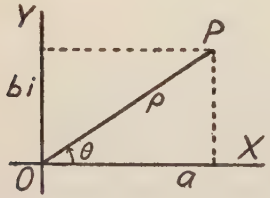


FIG. 88

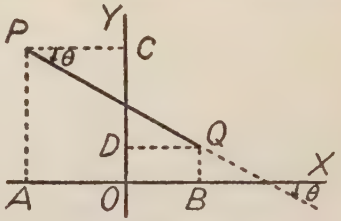


FIG. 89

and θ the angle it makes with the axis of reals. The segment PQ (Fig. 89) is an example in which a = AB is positive, b = CD, negative, ρ is the length PQ, and θ the angle CPQ, a fourth quadrant angle.

It is noted that if two line segments are parallel and of equal length and are similarly directed they represent the same complex number. The location of the line segment by which we represent a given number is not determined by the number itself, but may be chosen in relation to other features of the problem in hand.

We have here arbitrarily defined the geometrical representation for a complex number. That this representation is suitable, convenient and useful will be more apparent when we have considered the geometrical aspect of certain operations on complex numbers.

3. **Operations on complex numbers.** If we combine two numbers, as 3 + 2i, and − 5 + 2i, by **addition,** the result is −2i+ 4i. Geometrically 3 + 2i is represented by OP (Fig. 90), − 5 + 2i by PQ and the sum, −2 + 4i, by OQ. Reversing the order we might take −5 + 2i as OR and 3 + 2i as RQ and obtain the same result. The sum, then, ap-

pears as the diagonal of a parallelogram the sides of which represent the two given numbers. It is evident that this relation holds in any case of addition of complex numbers. The sum of several complex numbers may be found by placing end to end the corresponding line segments AB, BC, CD, DE, the sum appearing as the closing line segment, as AE. The geometrical process is identical with that in using the force polygon to find the resultant of a system of concurrent forces in a plane

Subtraction may always be accomplished by adding the negative of the number to be subtracted, but its geometrical aspect appears in figure 90 where the difference OQ − OP = PQ represents the difference $-2 + 4i - (3 + 2i) = -5 + 2i$.

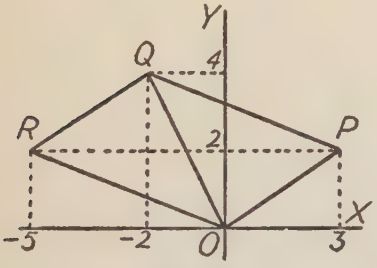


Fig. 90

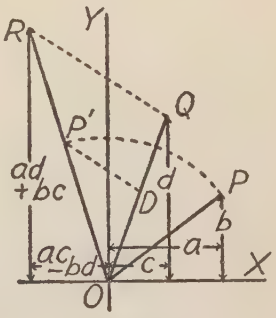


Fig. 91

Now let us illustrate **multiplication** by an example. Take

$$3 - 4i = 5e^{i\theta}\left(\sin\theta = -\frac{4}{5}, \cos\theta = \frac{3}{5}\right),$$

$$12 + 5i = 13e^{i\phi}\left(\sin\phi = \frac{5}{13}, \cos\phi = \frac{12}{13}\right).$$

Multiplying on the left and on the right we have

$$56 - 33i = 65e^{i(\theta+\phi)}.$$

This result is shown to be consistent by the relations

$$\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi = -\frac{4}{5}\cdot\frac{12}{13} + \frac{3}{5}\cdot\frac{5}{13} = -\frac{33}{65},$$

$$\cos(\theta + \phi) = \cos\theta\cos\phi - \sin\theta\sin\phi = \frac{3}{5}\cdot\frac{12}{13} + \frac{4}{5}\cdot\frac{5}{13} = \frac{56}{65}.$$

Now we may write any two numbers and their product in the three standard forms.

$$a + bi = \rho(\cos\theta + i\sin\theta) = \rho e^{i\theta}$$

$$c + di = \sigma(\cos\phi + i\sin\phi) = \sigma e^{i\phi}$$

$$(ac - bd) + (ad + bc)i = \rho\sigma[\cos(\theta + \phi) + i\sin(\theta + \phi)] = \rho\sigma e^{i(\theta+\phi)}.$$

This result may be obtained geometrically. Let the line segments OP and OQ (Fig. 91) represent the given numbers. Draw OR making the angle $\theta + \phi$ with OX. Take OP′ on OR of length ρ and take OD of unit length on OQ. Now, if QR is parallel to DP′, OR is the length of the required product.

We see then, that the product of two complex numbers has for modulus the product of the moduli of the numbers and for amplitude the sum of their amplitudes. We easily reach a similar conclusion for **division.** Since the dividend is the product of quotient and divisor, the modulus of the quotient is the modulus of the dividend divided by that of the divisor, and the amplitude of the quotient is that of the dividend diminished by that of the divisor. The geometrical aspect of division may be seen in Figure 91. Dividing OR by OQ we obtain OP. To get this result geometrically we take OD of unit length on OQ and take P′ on OR so that OP′ is parallel to QR, then lay off OP with length OP′ so that the angle POX is equal to the angle ROQ.

It is evident that multiplication and division are more easily accomplished when the numbers are expressed in the second or third standard form than when they are in the first standard form. If the modulus of a number is unity, that is, if the number may be written as $\cos\theta + i\sin\theta$ or $e^{i\theta}$, then multiplying any given number by it does not change the modulus of that number but rotates it through the angle θ. Thus, in figure 91, if Q coincides with D, the product of OP by OQ is OP′, having the same modulus as OP, but rotated through the angle ϕ. Similarly, division by a number whose modulus is unity merely rotates the dividend through an angle which is the negative of the amplitude of the divisor. For division by $e^{i\theta}$ is equivalent to multiplication by $e^{-i\theta}$.

In particular we have

$$i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = e^{i\frac{\pi}{2}},$$

and multiplication by i is equivalent to rotation through 90° counter

clockwise, while division by i is equivalent to multiplication by $-i$, and has the effect of rotation through 90° clockwise. Also we have

$$i^2 = -1 = \cos \pi + i \sin \pi = e^{i\pi},$$

and multiplying by i^2 has the effect of rotating through 180°, counter clockwise, while division by i^2 rotates through 180° clockwise. Either multiplication or division by -1 has the effect of reversing the direction of a number.

The effect of multiplying or dividing any number by a real positive number is merely to change its modulus, or length, leaving its amplitude unchanged. Thus in figure 91, if $\theta = 0$, OR would lie along OQ.

Regarding a complex number $\rho e^{i\theta}$ as an operator, to operate on other numbers by multiplication or division, it is convenient to think of its two factors as having distinct effects. The one factor ρ, having the effect of changing the length, is called a **tensor,** the other factor $e^{i\theta}$, having the effect of rotating through the angle θ, is called a **rotor.**

Now let us consider **involution** and **evolution** in relation to complex numbers. We evidently have

$$(a + bi)^k = \rho^k (\cos \theta + i \sin \theta)^k = \rho^k e^{ik\theta}. \tag{5}$$

Giving attention to this last equality we see that

$$(\cos \theta + i \sin \theta)^k = \cos k\theta + i \sin k\theta, \tag{6}$$

an important relation known as **DeMoivre's Theorem.** If we take k as a positive integer and expand the left member by the binomial theorem and equate reals to reals and imaginaries to imaginaries we obtain expressions for $\cos k\theta$ and $\sin k\theta$ in terms of $\cos \theta$ and $\sin \theta$.

Now relations (5) and (6) hold for all values of k, positive, negative, fractional. They therefore cover the process of finding roots as well as powers of a complex number. To see the full significance of these relations, however, we must note that the expression

$$\cos (\theta + 2n\pi) + i \sin (\theta + 2n\pi) = e^{i(\theta + 2n\pi)}, \tag{7}$$

where n is a positive or negative integer, has the same value as when $n = 0$. We have then, from (5),

$$(a + bi)^k = \rho^k e^{i(k\theta + 2nk\pi)}. \tag{8}$$

If k is a positive or negative integer this has but one value, found by putting $n = 0$, but if k is a rational fraction it will have several values. Let $k = \frac{p}{q}$, where p and q are positive integers having no common

factors, then the number will have q distinct values corresponding to $n = 0, 1, 2, \ldots, q - 1$. The values for $n = q, q + 1$, etc., will coincide with those for $n = 0, 1$, etc., respectively.

As an illustration let us find the value of

$$(1 + \sqrt{-3})^k = 2^k \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3}\right)^k.$$

First take $k = 5$, then the result is:

$$(1 + \sqrt{-3})^5 = 32 \left(\cos \frac{5\pi}{3} + i \sin \frac{5\pi}{3}\right) = 32 \left(\frac{1}{2} - \frac{1}{2}\sqrt{3}\, i\right)$$

$$= 16\,(1 - \sqrt{-3}).$$

Second, taking $k = \frac{1}{4}$ we have:

$$\sqrt[4]{1 + \sqrt{-3}} = 2^{\frac{1}{4}} \left[\cos\left(\frac{\pi}{12} + \frac{2n\pi}{4}\right) + i \sin\left(\frac{\pi}{12} + \frac{2n\pi}{4}\right)\right].$$

The four distinct values are

$$2^{\frac{1}{4}} (\cos 15° + i \sin 15°) \quad (n = 0)$$

$$2^{\frac{1}{4}} (\cos 105° + i \sin 105°) \quad (n = 1)$$

$$2^{\frac{1}{4}} (\cos 195° + i \sin 195°) \quad (n = 2)$$

$$2^{\frac{1}{4}} (\cos 285° + i \sin 285°) \quad (n = 3)$$

It should be noted that these four values have the same modulus, and that the amplitudes differ by multiples of 90°, so that their geometrical representations are equally spaced about the origin.

As further illustration, let us solve the equation $x^5 + 1 = 0$. Here we have

$$x = \sqrt[5]{-1} = (\cos \pi + i \sin \pi)^{\frac{1}{5}} = e^{\frac{i(2n+1)}{5}\pi}$$

Taking $n = 0, 1, 2, 3$, and 4 we have

$$x = \cos 36° + i \sin 36°,$$

$$\text{or} \quad \cos 108° + i \sin 108°,$$

$$\text{or} \quad \cos 180° + i \sin 180°,$$

$$\text{or} \quad \cos 252° + i \sin 252°,$$

$$\text{or} \quad \cos 324° + i \sin 324°.$$

These numbers are represented geometrically as radii (Fig. 92) of the unit circle, OA, OB, OC, OD, and OE. Only one root is real, $\cos 180° + i \sin 180° = -1$, which is represented by OC.

4. **Functions of a complex variable.** Consider $z = x + iy$, where x and y are variables whose values are real numbers. The quantity z is called a complex variable. By $w = f(z)$ we may denote a function of a complex variable. A study of such functions forms an important branch of mathematics, and it is worth while here to state a few features of such a study.

Just as in the discussion of functions of a real variable, in Chapter I, we regard a function as defined when a means is stated by which the value of the function w may be determined for each particular value that the independent variable z may take. And, just as in Chapter I, various means may be employed to determine the values of w, giving rise to different classes of functions.

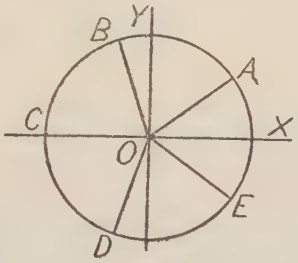


Fig. 92

If w is an **algebraic function** of z, then its value for any particular value of z may be found by the process discussed in Section 3 above. We have here, of course, the various types of algebraic functions; polynomials, rational fractions, irrational functions.

Next, let us define the **exponential function.** In Section 1 above we we defined $e^{i\theta}$ as a power series, and now we take as definition

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots \tag{9}$$

But if the usual properties of the exponential function are to hold we must have:

$$e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y). \tag{10}$$

By the definition this means that

$$1 + (x + iy) + \frac{(x + iy)^2}{2!} + \cdots$$

$$= \left(1 + x + \frac{x^2}{2!} + \cdots\right)\left[\left(1 - \frac{y^2}{2!} + \cdots\right) + i\left(y - \frac{y^3}{3!} + \cdots\right)\right].$$

This relation may be verified by addition and multiplication of series explained in Art. 49.

We define **logarithm** in the usual manner. We take $w = \log z$ as equivalent to $z = e^w$. But by (7) above, we have $e^w = e^{w+2in\pi}$ where n is any integer, therefore, if

$$z = \rho\,[\cos(\theta + 2n\pi) + i \sin(\theta + 2n\pi)] = e^{\log\rho}\, e^{i(\theta+2n\pi)}$$

we may write

$$\log z = \log \rho + i\,(\theta + 2n\pi). \tag{11}$$

Here $\log \rho$ is the natural logarithm of the positive real number ρ, the modulus of z.

Since

$$z = e^w = 10^w \log_{10} e$$

we have

$$\log_{10} z = w \log_{10} e = \log_{10} e \cdot \log z = \log_{10} \rho + i\,(\theta + 2n\pi) \log_{10} e. \tag{12}$$

By equations (11) and (12) we are able to write down the natural and common logarithms of any complex number. In particular, the amplitude θ for any negative real number is π, so that the quantity $\theta + 2n\pi$ becomes an odd multiple of π. In adding logarithms to find the logarithm of a product of several real factors, the result will involve an even or odd multiple of π according as the number of negative factors is odd or even. This justifies the usual method of dealing with negative factors in logarithmic computation.

Before defining the **trigonometric functions** we note that since $e^{i\theta} = \cos\theta + i \sin\theta$ and $e^{-i\theta} = \cos\theta - i \sin\theta$ we have

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}, \quad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}.$$

These are known as **Euler's expressions** for sine and cosine.

Now we may take as definitions the equations

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \quad \cos z = \frac{e^{iz} + e^{-iz}}{2} \tag{13}$$

or we may define these functions by power series in z of the same form as those derived for the real variable,

$$\begin{aligned} \sin z &= z - \frac{z^3}{3!} + \frac{z^5}{5!} + \ldots, \\ \cos z &= 1 - \frac{z^2}{3!} + \frac{z^4}{4!} + \ldots \end{aligned} \tag{14}$$

By applying (9) to (13) and performing the indicated operations on the resulting series, we may show that the two forms of definitions are equivalent. The tangent, cotangent, secant and cosecant of x, may now be defined as ratios and reciprocals of sine and cosine, according to the relations which hold in trigonometry.

For the **hyperbolic functions** we take as definitions

$$\sinh z = \frac{e^z - e^{-z}}{2}, \quad \cosh z = \frac{e^z + e^{-z}}{2}, \tag{15}$$

and the relations

$$\begin{aligned} \tanh z &= \frac{\sinh z}{\cosh z}, \quad \coth z = \frac{\cosh z}{\sinh z}, \\ \operatorname{csch} z &= \frac{1}{\sinh z}, \quad \operatorname{sech} z = \frac{1}{\cosh z}. \end{aligned} \tag{16}$$

From (13) and (15) we have

$$\begin{aligned} i \sin z &= \sinh (iz), \quad \cos z = \cosh (iz), \\ i \sinh z &= \sin (iz), \quad \cosh z = \cos (iz), \end{aligned} \tag{17}$$

and putting $z = x + iy$ in (13) and applying (15) and (17) we have

$$\begin{aligned} \sin (x + iy) &= \sin x \cosh y + i \cos x \sinh y, \\ \cos (x + iy) &= \cos x \cosh y - i \sin x \sinh y. \end{aligned} \tag{18}$$

Similarly we obtain

$$\begin{aligned} \sinh (x + iy) &= \sinh x \cos y + i \cosh x \sin y, \\ \cosh (x + iy) &= \cosh x \cos y + i \sinh x \sin y. \end{aligned} \tag{19}$$

From the definitions employed here it is not difficult to obtain all the standard formulas that are found in trigonometry, and similar formulas may be derived for the hyperbolic functions.

For the **graphical representation of functions** of a complex variable we note that such a function may be written

$$w = f(z) = u + iv,$$

where, since $z = x + iy$, the real quantities u and v are functions of x and y. Now using the plan of representing complex numbers that we explained in section 2 above we may represent any value of z in one plane and the corresponding value of w in another plane. Given any value of $z = x + iy$, we represent z by OP (Fig. 93) where x and y are the coordinates of P. Similarly w is represented by QR, the coordinates of R being u and v, which are computed as functions of the known x and y. The variable z may take values corresponding to any position of P in the z-plane, and for each value of z there will be a definite value of w and a definite position of R in the w-plane. If P

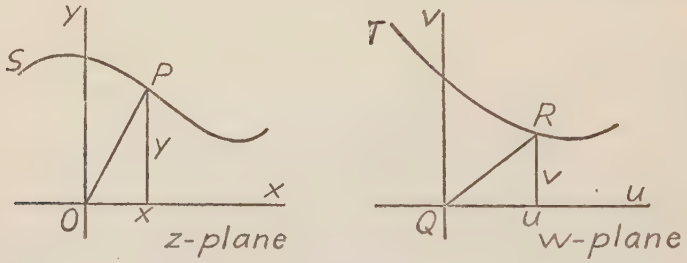


Fig. 93

describes a locus PS, then R will describe some locus, as RT. To illustrate take

$$w = z^2 = (x + iy)^2 = x^2 - y^2 + 2ixy.$$

Here we have $u = x^2 - y^2$, $v = 2xy$. Let z describe the line whose equation is $y = mx$, then the path of w may be expressed by the parametric equations $u = (1 - m^2)x^2$, $v = 2mx^2$, or by eliminating x,

$$v = \frac{2m}{1 - m^2} u.$$

If z describes a circle $x^2 + y^2 = r^2$ then we have $u^2 + v^2 = r^4$, also a circle. If z describes a vertical line $x = a$, then $u = a^2 - y^2$, $v = 2ay$, or $v^2 = -4a^2(u - a^2)$, a parabola. Corresponding to any net of curves in the z-plane there will be a net of curves in the w-plane. The z-plane is said to be mapped on the w-plane by the particular

function in hand. In Figure 94 the curves given by the above equations are drawn for several values of m, r and a, also the curve for y = b = 1.5. Since two values of z correspond to one value of w, the whole w-plane is required to map the first and second quadrants of the z-plane. Similarly the third and fourth quadrants may be mapped on another w-plane. Four shaded regions of the z-plane are shown mapped on the w-plane by corresponding shading lines. Corresponding points are given the same letters and corresponding lines or curves the same numbers in the two planes.

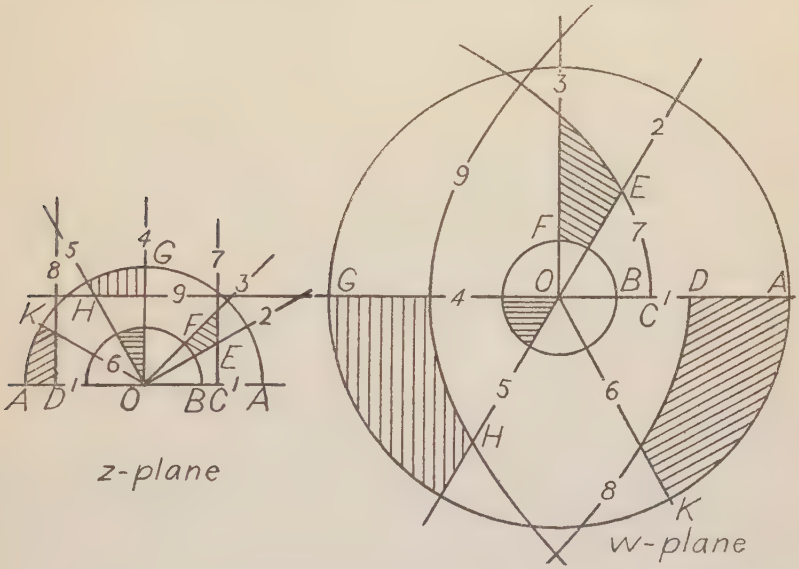


FIG. 94

In a similar way any given function will furnish a means of mapping the z-plane on the w-plane, the character of the map being different for each function.

It lies beyond the scope of the present work to discuss further the properties of functions of complex variables. But it may be stated that the processes of taking derivatives and integrating as we have applied them to the real variables x and y are applicable to the simple functions defined in the present article. The necessary formulas are obtained by putting z and w for x and y in the corresponding formulas in the Chapters III and VI. Of course the graphical interpretations of these processes given in those early chapters are inadequate here.

EXERCISES XLVII

1. Write each of the following numbers in the three standard forms for complex numbers.

(a) $5 + 12\sqrt{-1}$; (b) $2 - \sqrt{-21}$; (c) $\sqrt{2}\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right)$;

(d) $e^{i\frac{\pi}{3}}$; (e) $\frac{3+\sqrt{-2}}{4-\sqrt{-7}}$; (f) $3e^{2i}$.

2. Make combinations of complex numbers as follows:

(a) $(5 - 3i) + (8 + 7i)$; (b) $(4 + 9i)(3 - 2i)$; (c) $e^{i\frac{\pi}{4}} - e^{i\frac{\pi}{6}}$;

(d) $3e^{-i\pi} \cdot \frac{1}{5}e^{i\frac{\pi}{3}}$; (e) Divide $2 + \sqrt{-12}$ by $5e^{i\frac{\pi}{6}}$.

3. Obtain the results called for in 2, by graphical methods, using the geometrical representation of each number involved.

4. Obtain the following powers by the binomial theorem, then by use of DeMoivre's theorem, checking results by comparison:

(a) $(3 + 4i)^3$; (b) $\left(2e^{i\frac{\pi}{3}}\right)^5$; (c) $\left(\frac{5}{13} - \frac{12}{13i}\right)^4$.

5. Find the indicated roots, showing the multiple values geometrically:

(a) $\sqrt[3]{1}$; (b) $\sqrt[6]{-1}$; (c) $\sqrt[4]{7-5i}$; (d) $\sqrt[5]{32(\cos 75° + i\sin 75°)}$.

6. Solve the following equations:

(a) $x^3 + 1 = 0$; (b) $x^4 - 1 = 0$; (c) $x^5 + 1 = 0$;

(d) $x^4 + 7 = 0$; (e) $x^5 - 13 = 0$; (f) $x^6 + 4x^3 + 8 = 0$.

7. Observing that $x^5 + a^5 = (x^4 - ax^3 + a^2x^2 - a^3x + a^4)(x + a)$, solve the equation

$$x^4 - 2x^3 + 4x^2 - 8 + 16 = 0.$$

8. Express $\cos 5\theta$ in terms of $\sin\theta$ and $\cos\theta$.

9. If $z = x + iy = 2 + i\frac{\pi}{3}$, express each of the following functions of z as a complex number in the form $a + bi$.

(a) e^{2z}; (b) $\log(z - 1)$; (c) $\sin\frac{1}{2}z$; (d) $\tan 2z$;

(e) $\cosh 3z$; (f) $e^{-\frac{1}{2}z}$; (g) $\sec iz$.

10. Prove the following relations:

(a) $\cosh^2 z - \sinh^2 z = 1$; (b) $1 - \tanh^2 z = \operatorname{sech}^2 z$;
(c) $\sinh(z_1 + z_2) = \sinh z_1 \cosh z_2 + \cosh z_1 \sinh z_2$.

11. Show that if $w = \cosh z$, then $z = \cosh^{-1} w = \log(w + \sqrt{w^2 + 1})$.

12. Show that if $w = \sin z$, then $z = \sin^{-1} w = -i \log(iw + \sqrt{1 - w^2})$.

13. If $w = z^2 - 2z$, what path must z describe in the z-plane in order that in the w-plane the path of w shall be the line $u = v$? What path does w describe when z takes the path $x = y$?

14. If $w = \cos z$, show that the following loci for z give straight line loci for w.

(a) $x = \frac{\pi}{2}$; (b) $x = 0$; (c) $y = 0$; (d) $\cot x = \tanh y$.

15. If $w = \sinh z$, write the equations of four loci in the z-plane which map as straight lines in the w-plane.

16. If $w = e^z$, show on the w-plane the map of the square in the z-plane bounded by the lines $x = 0$, $x = 1$, $y = 0$, $y = 1$.

17. Use logarithms to obtain the product

$$(-3.27)\,(\sqrt{-742})\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right)\left(179e^{i\frac{\pi}{5}}\right).$$

Art. 51. Simple Harmonic Functions

1. **Standard forms.** The student is familiar with the fact that all trigonometric functions are periodic, repeating their values in cycles as the independent variable increases. The increment in the independent variable that carries the function through a complete cycle of values is called the **period** of the function. In particular, a function expressible in the form

$$y = c \sin(kx + \phi), \tag{1}$$

where c, k and ϕ are constants, is a **simple harmonic function.** The coefficient c is the **amplitude,** ϕ is the **phase angle,** and k determines the period, since a complete cycle of values of y correspond to a change of kx through 2π, so that the period is $\frac{2\pi}{k}$. The number of cycles per unit change in x is $\frac{k}{2\pi}$, and this is called the **frequency.** It is evident that we may write $y = \cos\left(kx + \phi - \frac{\pi}{2}\right)$, so that a harmonic func-

tion may be expressed by either a sine or a cosine. Also we may write

$$y = a \cos kx + b \sin kx \tag{2}$$

where $a = c \sin \phi$ and $b = c \cos \phi$. Forms (1) and (2) are standard forms for writing a simple harmonic function.

By use of equations (13) of Art. 50 we have

$$c \sin (kx + \phi) = \frac{e}{2i} \left[e^{i(kx+\phi)} - e^{-i(kx+\phi)}\right]$$

$$= \frac{c}{2i} e^{i\phi} \cdot e^{kix} - \frac{c}{2i} e^{-i\phi} \cdot e^{-kix}.$$

Also,

$$a \cos kx = \frac{a}{2} (e^{kix} + e^{-kix}),$$

$$b \sin kx = \frac{b}{2i} (e^{kix} - e^{-kix}),$$

therefore, noting that the reciprocal of i is $-i$,

$$y = \frac{a - bi}{2} e^{kix} + \frac{a + bi}{2} e^{-kix}.$$

Now, assuming that (1) and (2) represent the same function, we may write

$$A = \frac{c}{2i} e^{i\phi} = \frac{a - bi}{2}, \quad B = \frac{c}{2i} e^{-i\phi} = \frac{a + bi}{2},$$

and obtain the result

$$y = Ae^{kix} + Be^{-kix} \tag{3}$$

This is a third standard form for writing a harmonic function, but in order that this form be real it is necessary that A and B shall be conjugate numbers, that is, their real parts must be equal and their pure imaginary parts equal in magnitude but opposite in sign.

Any simple harmonic function may be written in any one of the three forms, and involves three constants; c, k and ϕ in (1); a, b and k in (2); and A, B and k in (3). Harmonic functions are often arrived at in the form (3) as a result of mathematical processes, such as the

solution of differential equations, but are usually transformed to form (2) or form (1) when it is desired to discuss the function.

2. **Combinations of simple harmonic functions.** Of course any two functions of a variable may be added to obtain another function, but in the case of harmonic functions special significance is attached to addition of functions having the same period but differing in amplitude and phase. Let us express two functions as

$$y_1 = a_1 \cos kx + b_1 \sin kx,$$

$$y_2 = a_2 \cos kx + b_2 \sin kx,$$

then their sum is

$$y = y_1 + y_2 = (a_1 + a_2) \cos kx + (b_1 + b_2) \sin kx.$$

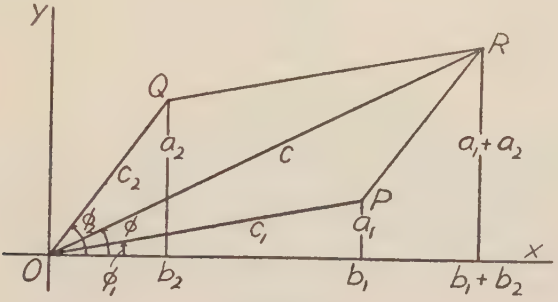


FIG. 95

Now associate this with the process of adding two complex numbers,

$$(b_1 + a_1 i) + (b_2 + a_2 i) = (b_1 + b_2) + (a_1 + a_2)i,$$

and with the geometrical process corresponding to this addition, as treated in Section 2 of Art. 50. In Figure 95 the line segments OP, OQ, and OR are associated with the complex numbers

$$\begin{aligned} b_1 + a_1 i &= c_1 e^{i\phi_1}, \\ b_2 + a_2 i &= c_2 e^{i\phi_2}, \\ (b_1 + b_2) + (a_1 + a_2)\, i &= c e^{i\phi}. \end{aligned} \tag{4}$$

If we associate them with the three harmonic functions

$$\begin{aligned} a_1 \cos kx + b_1 \sin kx &= c_1 \sin (kx + \phi_1), \\ a_2 \cos kx + b_2 \sin kx &= c_2 \sin (kx + \phi_2), \\ (a_1 + a_2) \cos kx + (b_1 + b_2) \sin kx &= c \sin (kx + \phi), \end{aligned} \tag{5}$$

then it is evident that the geometrical process of combining vectors may be substituted for the process of adding the associated harmonic functions. It is convenient, then to take the expressions (4) as **symbolic representations** for the harmonic functions (5). Any number of harmonic functions may be combined by addition and subtraction by adding and subtracting the complex numbers which represent them symbolically. Since $-e^{i\phi} = e^{i(\phi + \pi)}$, subtraction of a harmonic functions is equivalent to addition of one in opposite phase.

In applying the symbolic treatment to harmonic functions it is well to note that the period or frequency is left out of consideration, and only functions having the same frequency can be treated together by this method. The symbolic representation (4) must not be confused with the exponential form (3). The relation of the symbolic representation to the function itself may be seen more clearly in Figure

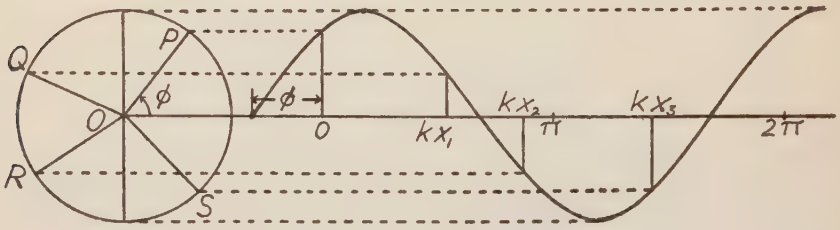


FIG. 96

96. Let OP be the graphical representation of the number $ce^{i\phi}$ which is the symbolic representation of the harmonic function

$$c \sin (kx + \phi).$$

The value of the function for any value x_1 of x is found by taking the angle POQ equal to kx_1 radians, then projecting OQ on the vertical. We obtain a graph of the function by laying of a horizontal scale for kx in radians and erecting ordinates equal to the value of the function for various values of x. The construction is shown for $kx = 0$, $kx_1 = POQ$, $kx_2 = POR$, and for $kx_3 = POS$, also for $kx + \phi = 0, \frac{\pi}{2}, \frac{3\pi}{2}$.

The full geometrical significance of addition of harmonic functions is shown in Figure 97. Let OP and OQ represent symbolically two functions, and OR their sum. For any x take angles POP′, QOQ′ and ROR′ each equal to kx radians and project P′, Q′ and R′ onto the

vertical at D, E and F, then ordinates OD and OE are the values of the two given functions and OF that of their sum. As x varies the parallelogram OP′R′Q′ revolves about O, remaining constant in size and shape. The ordinates OD and OE vary, going through complete cycles while kx changes through 2π, but their sum is always equal to the variable ordinate OF.

3. **Polyphase relations.** We have seen above how two or more harmonic functions having the same period may be combined by the parallelogram law, just as we deal with the sum of complex numbers or the resultant of concurrent forces in a plane. Now let us consider n harmonic functions of common amplitude and frequency having

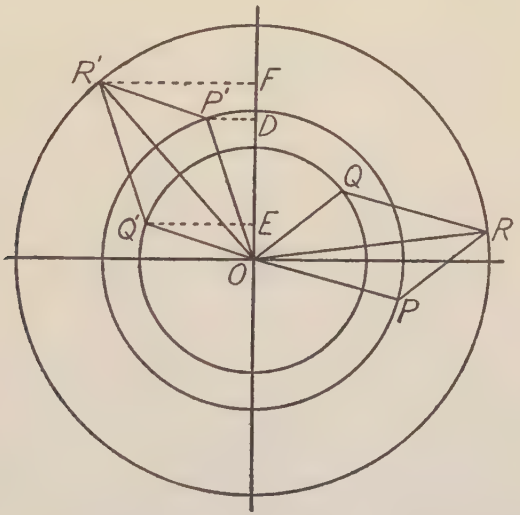


Fig. 97

phase angles differing by successive multiples of $\frac{2\pi}{n}$. As an example of such a set of functions take

$$c \sin (kx + \phi),\ c \sin \left(kx + \phi + \frac{2\pi}{n}\right),\ c \sin \left(kx + \phi + \frac{4\pi}{n}\right),$$

$$\ldots\ .\ .\ .\ .\ c \sin \left(kx + \phi + \frac{n-1}{n} 2\pi\right). \qquad (6)$$

The symbolic representations of these functions appear geometrically as lines of length c, all radii of the same circle, the angular differences

being $\frac{2\pi}{n}$. The sum of such a set is evidently zero, so that we may write the equation,

$$\sum_{r=0}^{r=n-1} \sin\left(kx + \phi + \frac{2r\pi}{n}\right) = 0, \tag{7}$$

and this result is true for all values of k, x, ϕ and n. In Figure 98 the lines OP_1 to OP_5 represent the five functions when $n = 5$.

Again, let us consider the set of functions

$$c \sin(kx + \phi), c \sin\left(kx + \phi + \frac{2m\pi}{n}\right), \ldots .$$

$$\ldots . , c \sin\left(kx + \phi + \frac{n-1}{n} 2m\pi\right), \tag{8}$$

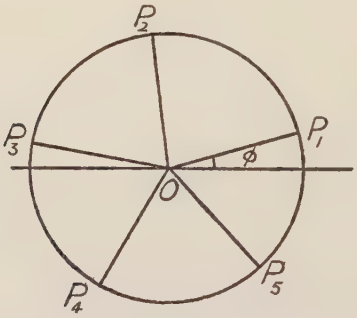


Fig. 98

where m is an integer. Here the phase differences are $\frac{2m\pi}{n}$. Let the highest common factor of m and n be p, and let $m = m'p$, $n = n'p$. The n phase angles differ from ϕ by angles of the form $2m'\pi M$, where the n values of M may be arranged in p rows and n′ columns as follows:

$$\begin{array}{llll}
0, & \frac{1}{n'}, & \frac{2}{n'}, \ldots . , & \frac{n'-1}{n'}, \\
1, & 1+\frac{1}{n'}, & 1+\frac{2}{n'}, \ldots . , & 1+\frac{n'-1}{n'}, \\
2, & 2+\frac{1}{n'}, & 2+\frac{2}{n'}, \ldots . , & 2+\frac{n'-1}{n'}, \\
\ldots & \ldots & \ldots & \ldots \\
p-1, & p-1+\frac{1}{n'}, & p-1+\frac{2}{n'}, \ldots . , & p-1+\frac{n'-1}{n'}.
\end{array}$$

Now since no two numbers $\frac{2m'\pi r_1}{n'}$ and $\frac{2m'\pi r_2}{n'}$ can differ by a multiple of 2π unless $r_2 - r_1$ is a multiple of n', the functions corresponding to the first row above are all distinct, and if $n' > 1$, they have a sum equal to zero. The phase angles given by the second row exceed those above them by $2m'\pi$, therefore the functions corresponding to them coincide with those corresponding to the first row, and similarly for all the rows. Therefore, if $n' > 1$, that is, **if m is not a multiple of n**, we have

$$\sum_{r=0}^{r=n-1} \sin\left(kx + \phi + \frac{2mr\pi}{n}\right) = 0. \tag{9}$$

If $n' = 1$, then $p = n$ and the n values of M are given by the first column above, the n functions coincide, and therefore, **if m is a multiple of n,**

$$\sum_{r=0}^{r=n-1} \sin\left(kx + \phi + \frac{2mr\pi}{n}\right) = n \sin(kx + \phi) \tag{10}$$

In particular, let $n = 8$ and let $m = 4$, then $p = 4$, $m' = 1$, $n' = 2$. The functions are:

$$\begin{aligned} &c \sin(kx + \phi), && c \sin(kx + \phi + \pi),\\ &c \sin(kx + \phi + 2\pi), && c \sin(kx + \phi + 3\pi),\\ &c \sin(kx + \phi + 4\pi), && c \sin(kx + \phi + 5\pi),\\ &c \sin(kx + \phi + 6\pi), && c \sin(kx + \phi + 7\pi). \end{aligned}$$

The four functions in each column are all equal, and the two functions in each row differ only in sign, so that the sum of the 8 functions is zero.

Again, take $n = 9$ and $m = 6$, then $p = 3$, $m' = 2$ and $n' = 3$. Now the phase angles differ from ϕ by the following 9 values:

$$\begin{aligned} &0, && \frac{4\pi}{3}, && \frac{8\pi}{3},\\ &4\pi, && 4\pi + \frac{4\pi}{3}, && 4\pi + \frac{8\pi}{3},\\ &8\pi, && 8\pi + \frac{4\pi}{3}, && 8\pi + \frac{8\pi}{3}. \end{aligned}$$

Each of the three rows will give the same set of functions, and the set of three functions corresponding to one row evidently has a sum equal to zero.

As a final example take $n = 5$, $m = 4$, then $p = 1$, $m' = 4$, and $n' = 5$. The phase angles differ from ϕ by the following values,

$$0, \quad \frac{8\pi}{5}, \quad \frac{16\pi}{5}, \quad \frac{24\pi}{5}, \quad \frac{32\pi}{5},$$

and these values are equivalent to

$$0, \quad -\frac{2\pi}{5}, \quad -\frac{4\pi}{5}, \quad -\frac{6\pi}{5}, \quad -\frac{8\pi}{5},$$

therefore the five functions are represented as in figure 98, but in reverse order.

EXERCISES XLVIII

1. Express each of the following functions in the three standard forms and state the value of the amplitude, period, and phase angle:

(a) $4 \sin 6x - 3 \cos 6x$; (b) $142 \sin\left(4x + \frac{\pi}{4}\right)$;

(c) $5e^{3ix} + 5e^{-3ix}$; (d) $2i(e^{ix} - e^{-ix})$;

(e) $(7 - 5i)e^{8ix} + (7 + 5i)e^{-8ix}$.

2. Show that the following are simple harmonic functions:

(a) $4 \cos^2 x - 2$; (b) $\sqrt{1 - \cos 6x}$; (c) $3 \sin 2x \cos 2x$;

(d) $\frac{1}{2} - \sin^2\left(\frac{x + \pi}{3}\right)$; (e) $7\left[\sin^2\left(\frac{x}{3}\right) - \cos^2\left(\frac{x}{3}\right)\right]$.

3. Write a harmonic function of the time t in seconds which shall have 10 cycles per second and which shall have a value 2 when $t = 0$ and a value 5 when $t = \frac{1}{40}$.

4. Find the sum of the harmonic functions $S_1 = 10 \sin 300t$ and $S_2 = 10 \sin\left(300t + \frac{2\pi}{3}\right)$.

5. A three-phase electric transmission line carries three currents, the equations of which are $i_1 = I \sin \omega t$, $i_2 = I \sin\left(\omega t + \frac{2\pi}{3}\right)$ and $i_3 = I \sin\left(\omega t + \frac{4\pi}{3}\right)$. Show that the sum of the three currents at any instant is zero.

6. The centrifugal force on a rotating crank has a magnitude $M\omega^2$, where M is the mass of the rotating parts multiplied by the distance of its center of gravity from the shaft axis and ω is the angular speed in radians per second. The components of this force along two rectangular axes are $M\omega^2 \sin\theta$ and $M\omega^2 \cos\theta$, where θ is the angular displacement of the crank from one of these axes. Show that the forces on n equal cranks arranged at angular intervals of $\frac{2\pi}{n}$ are balanced. Draw the diagram for $n = 4$.

7. Consider example 6, as applying to a reciprocating engine, where $\theta = o$ at the beginning of the out stroke. Now the moment of the force on a crank about a line parallel to the stroke through the shaft axis is $aM\omega^2 \sin\theta$, where a is distance from the line to the plane of the crank. Show that if the cranks are equal, arranged at equal angular intervals and spaced at equal intervals along the shaft, then the moments cannot balance. Draw diagrams for $n = 3$ and $n = 5$, taking the moment axis in the plane of the first crank.

8. The force due to the action of reciprocating parts at a crank is approximately

$$F = \frac{W}{g} r\omega^2 \cos\theta + \frac{W}{g}\frac{r^2}{l}\omega^2 \cos 2\theta,$$

where $\frac{W}{g}$ is the mass of reciprocating parts, r is the crank radius, l is the length of connecting rod and other notations are as in examples 6 and 7. The first term is called the primary force and the other the secondary force. If the arrangement of cranks is regular as in 6 and 7, show that primary forces are balanced for $n > 1$, and secondary forces are balanced for $n > 2$, but not for $n = 2$.

Art. 52. Analysis of Periodic Functions

1. **Fourier series.** The name Fourier is attached to series of the form

$$c_o + c_1 \sin(x + \phi_1) + \ldots + c_n \sin(nx + \phi_n) \ldots \quad (1)$$

Each term after the first is a simple harmonic function, and by the discussion of the previous article we see that an equivalent form is

$$a_o + a_1 \cos x + a_2 \cos 2x + \ldots + a_n \cos nx + \ldots$$
$$+ b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx + \ldots \quad (2)$$

where $a_o = c_o$, $a_n = c_n \sin \phi_n$ and $b_n = c_n \cos \phi_n$. If such a series is convergent for all values of x then it represents a function of x, say f(x), and this function has the period 2π; for c_o is constant; the next term,

$$c_1 \sin (x + \phi_1) = a_1 \cos x + b_1 \sin x,$$

known as the **first harmonic** of f(x), completes a cycle of values when x changes through 2π; and any other term,

$$c_n \sin (nx + \phi_n) = a_n \cos nx + b_n \sin nx,$$

known as the **nth harmonic** of f(x), completes n cycles of values in the same period.

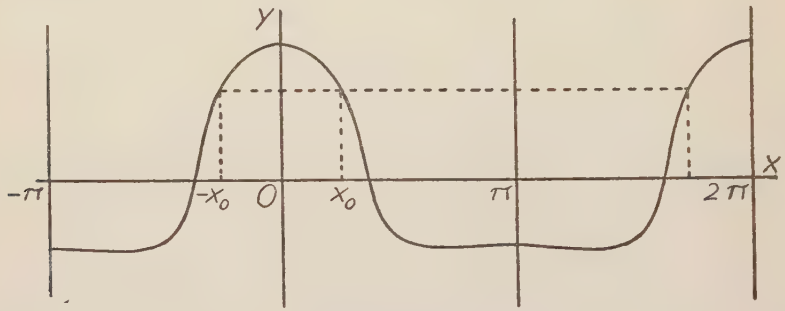


FIG. 99

Let us give attention to certain special features such a function may possess. First let us suppose that the series (2) consists of **cosine terms only,** so that

$$f(x) = a_1 \cos x + a_2 \cos 2x + \ldots + a_n \cos nx + \ldots . \quad (3)$$

In this case every term has the same value at x_o as at $-x_o$ or at $2\pi - x_o$, so that the graph is symmetrical with respect to the lines $x = o$, $x = \pi$, $x = 2\pi$, etc. (Fig. 99).

Again, suppose series (2) consists of **sine terms only,** so that

$$f(x) = b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx + \ldots \quad (4)$$

In this case the value of each term at x_o is the negative of its value at $-x_o$ or $2\pi - x_o$, so that the graph is symmetrical with respect to the points on the x-axis where x is 0, π, 2π, etc. (Fig. 100). Further, the graph will pass through these points of symmetry, since every term is zero at any multiple of π.

It is evident that a function containing any sine term could not have the kind of symmetry shown in Figure 99, and that a function containing a cosine term could not have the kind of symmetry shown in Figure 100.

2. **Values of coefficients.** Let series (2) represent a function $f(x)$, and suppose that the function may be integrated from $x = 0$ to $x = 2\pi$, and that the series may be integrated term by term on the same interval, then we have

$$\int_0^{2\pi} f(x)\,dx = \int_0^{2\pi} a_o\,dx + R_o,$$

where R_o consists of such terms as

$$\int_0^{2\pi} a_n \cos nx\,dx, \quad \int_0^{2\pi} b_n \sin nx\,dx,$$

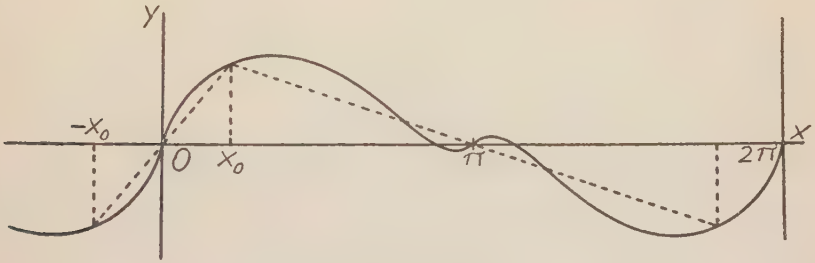


FIG. 100

each of which is evidently zero. We have, therefore

$$a_o = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx. \tag{5}$$

Again, let us multiply the function and the series each by $\cos mx$ and integrate, then

$$\int_0^{2\pi} f(x)\cos mx\,dx = \int_0^{2\pi} a_o \cos mx\,dx + R_m = R_m$$

where R_m consists of terms of the two forms

$$\int_0^{2\pi} a_n \cos nx \cos mx\,dx, \quad \int_0^{2\pi} b_n \sin nx \cos mx\,dx.$$

Now we have the trigonometric identity

$$\cos nx \cos mx = \frac{1}{2}\cos (m + n)x + \frac{1}{2}\cos (m - n)x,$$

and if $m \neq n$ these terms are simple harmonic functions and the integral is zero. If $m = n$ the last term is constant, $\frac{1}{2}$, and the integral has the value πa_m. Also we have the identity

$$\sin nx \cos mx = \frac{1}{2} \sin (m + n)x - \frac{1}{2} \sin (m - n)x,$$

and the integrals are zero for all values of n, since the last term is zero if $m = n$. We have, then, $R_m = \pi a_m$, and

$$a_m = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos mx \, dx. \tag{6}$$

Finally, let us multiply the function and the series by sin mx and integrate. We shall have

$$\int_0^{2\pi} f(x) \sin mx \, dx = \int_0^{2\pi} a_0 \sin mx \, dx + S_m = S_m,$$

where S_m consists of terms of the two forms

$$\int_0^{2\pi} a_n \cos nx \sin mx \, dx, \quad \int_0^{2\pi} b_n \sin nx \sin mx \, dx.$$

The first form here is zero for all values of n, just as the second form in the previous case. For the second form we have

$$\sin nx \sin mx = \frac{1}{2} \cos (m + n)x + \frac{1}{2} \cos (m - n)x,$$

so that the integral is zero when $m \neq n$, but gives πb_m when $m = n$. Therefore, $S_m = \pi b_m$
and

$$b_m = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin mx \, dx. \tag{7}$$

If in (6) and (7) we put m successively equal to 1, 2, 3, etc., we have a means for determining the coefficients in a Fourier series which shall represent the function f(x), assuming that all the processes involved are valid.

3. **Conditions of application.** It lies beyond the scope of the present work to establish the conditions under which a given function may be represented by a Fourier series, but we may state conditions sufficient for such representation based on the authority of more extended works on this subject.

(a) **Any continuous periodic function which has a finite number of maxima and minima in each cycle may be represented by a Fourier series.** If the function has a period T in terms of the independent variable x, then we may write $\frac{2\pi x}{T} = x'$, then in terms of the new variable x′ the period will be 2π, so that we may always take the series in the form (2), with a possibility of reduction to (3) or (4) by symmetry. The equations (5), (6) and (7) determine the constants uniquely.

(b) A function having the character stated in (a) above except for **finite discontinuities at a finite number of points** may be represented by a Fourier series, provided that the value of the function at a point of discontinuity, x_o, be taken as half the sum of the two unilateral limits.

$$\lim_{x \to x_o -} f(x), \qquad \lim_{x \to x_o +} f(x)$$

(c) Any function may be represented by a Fourier series **on a limited interval,** say from x = a, to x = b, if the periodic function, one complete cycle of which agrees with the given function on the stated interval, is found to fulfill the conditions of (a) or (b) above. If we write

$$x' = \frac{2\pi x - 2a\pi}{b - a}$$

then x′ = 0 when x = a and x′ = 2π when x = b, so that, in terms of x′, we have to deal with the interval from 0 to 2π. Whatever the function of x′ may be on this interval, we may form a periodic function by repeating this same cycle of values from 2π to 4π, and, in general, from $2n\pi$ to $2(n + 1)\pi$.

A Fourier series in form (1) is dominated by the series of positive constants

$$|c_o| + |c_1| + |c_2| + \ldots + |c_n| + \ldots,$$

therefore, if this series of constants converges the series must represent a continuous function, by the reasoning of Art. 49. By the same article, then, we are assured that we may integrate the function represented by integrating the series term by term. Further, if on differentiation of a Fourier series term by term the series of absolute values of coefficients of the new series is found to converge, then the

new series represents the derivative of the function represented by the given series.

But in the case of Fourier series we may go beyond the general results reached in Art. 49, and state that whenever a Fourier series represents a function under the conditions of (a), (b) or (c) the process of integration term by term yields valid results for the function represented. On the other hand, the results of differentiation term by term must not be accepted in the absence of adequate tests of their validity. The convergence of the series of coefficients of the derived series is a sufficient test, but not a necessary condition, since the derived function may well be discontinuous.

4. **Methods for evaluation of constants.** The function that we may wish to develop in a Fourier series may be represented in various ways; by an equation or formula, by a graph drawn mechanically or otherwise, by tabulated values determined experimentally or otherwise. The practical method for finding the values of the coefficients must necessarily depend on the character of the function and the way it is presented.

Fig. 101

(a) If the function is given by an equation or formula it may be convenient to do the actual integration indicated in equations (5), (6) and (7). As an example let it be required to represent the function $y = x$ on the interval from $x = 0$, to $x = 1$. Take $x' = 2\pi x$, then $y = \frac{x'}{2\pi}$ and we have to deal with the interval from $x' = 0$ to $x' = 2\pi$. Dropping the accent temporarily, we have by (5)

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} \frac{x\,dx}{2\pi} = \frac{1}{2}.$$

Now the function $y - a_0 = \frac{x}{2\pi} - \frac{1}{2}$ from 0 to 2π is one cycle of a periodic function (Fig. 101) which has the type of symmetry of (4), and therefore must contain sine terms only. In fact, if we should

apply (6) should find $a_m = 0$ for all values of m. But applying (7) we have

$$b_m = \frac{1}{\pi}\int_0^{2\pi} \frac{x}{2\pi} \sin mx\, dx$$

$$= \frac{1}{2\pi^2 m^2}\left[\sin mx - mx \cos mx\right]_0^{2\pi} = -\frac{1}{\pi m}.$$

The development is, restoring the accent,

$$\frac{x'}{2\pi} = \frac{1}{2} - \frac{1}{\pi}\left(\sin x' + \frac{1}{2}\sin 2x' + \ldots + \frac{1}{n}\sin n\,x' + \ldots\right),$$

and therefore in terms of the original variable we have

$$x = \frac{1}{2} - \frac{1}{\pi}\left(\sin 2\pi x + \frac{1}{2}\sin 4\pi x + \ldots + \frac{1}{n}\sin 2n\,\pi x + \ldots\right).$$

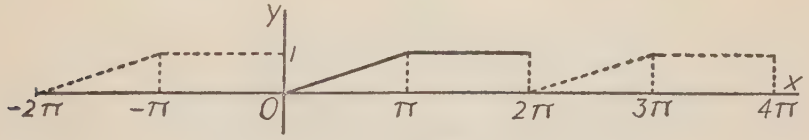


Fig. 102

This is a true equation for x between 0 and 1. But the periodic function on the right is discontinuous at integral values, and at these points it has the value $\frac{1}{2}$, which is half the sum of the two values that x takes at the ends of the interval, 0 and 1.

As another example, take $y = \frac{x}{\pi}$ from $x = 0$ to $x = \pi$, and take $y = 1$ from $x = \pi$ to $x = 2\pi$ (Fig. 102). From equation (5) we have

$$a_0 = \frac{1}{2\pi}\left[\int_0^{\pi} \frac{x}{\pi}\, dx + \int_{\pi}^{2\pi} dx\right] = \frac{3}{4}.$$

From equation (6) we obtain

$$a_m = \frac{1}{\pi}\left[\int_0^{\pi} \frac{x}{\pi} \cos mx\, dx + \int_{\pi}^{2\pi} \cos mx\, dx.\right]$$

$$= \frac{1}{(\pi m)^2}\left[mx \sin mx + \cos mx\right]_0^{\pi} + \frac{1}{\pi m}\left[\sin mx\right]_{\pi}^{2\pi}.$$

This gives $a_m = 0$ if m is even, and $a_m = -\frac{2}{(\pi m)^2}$ if m is odd. From (7) we have

$$b_m = \frac{1}{\pi}\left[\int_0^{\pi} \frac{x}{\pi} \sin mx\, dx + \int_{\pi}^{2\pi} \sin mx\, dx\right]$$

$$= \frac{1}{(\pi m)^2}\left[\sin mx - mx \cos mx\right]_0^{\pi} - \frac{1}{\pi m}\left[\cos mx\right]_{\pi}^{2\pi},$$

and this gives $b_m = -\frac{1}{\pi m}$ if m is even, and $b_m = -\frac{1}{\pi m}$ if m is odd.

Collecting results we have

$$y = \frac{3}{4} - \frac{2}{\pi^2}\left[\cos x + \frac{1}{9}\cos 3x + \frac{1}{25}\cos 5x + \ldots\right]$$

$$- \frac{1}{\pi}\left[\sin x + \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x + \frac{1}{4}\sin 4x + \ldots\right].$$

(b) Certain cases arise in which the function is given by an equation but in which it is impossible or inexpedient to perform the integrations indicated in (5), (6) and (7). In such cases it may be convenient to derive the development in a Fourier series by other means, at least to the extent of approximating the coefficients of the first few harmonics. If the function can be expressed in terms of a sine or a cosine in such manner as to yield a power series in powers of the sine or cosine, then these powers may be expressed in terms of multiple angles and the result arranged as a Fourier series.

To express powers of cos x in terms of multiple angles we have the identity

$$\cos^n x = \frac{1}{2^{n-1}}\left[\cos nx + n\cos(n-2)x + \frac{n(n-1)}{1\cdot 2}\cos(n-4)x + \ldots + R_n\right]. \qquad (8)$$

The coefficients of the expression in brackets are formed as binomial coefficients. If n is even the number of terms is $\frac{n}{2} + 1$, the last term, R_n being constant, equal to half of the coefficient of the middle term in the expansion of $(a + b)^n$. If n is odd the number of terms is $\frac{n+1}{2}$ and there is no constant term.

Equation (8) is easily derived by the use of Euler's expression for the cosine, by which we have

$$\cos^n x = \left(\frac{e^{ix} + e^{-ix}}{2}\right)^n = \frac{1}{2^n}\left[e^{inx} + ne^{i(n-2)x} + \frac{n(n-1)}{1\cdot 2}e^{i(n-4)x}\right.$$

$$\left. + \ldots\ldots + ne^{-i(n-2)x} + e^{-inx}\right].$$

In the brackets the first and last terms combine to give 2 cos nx, the second and next to last to give 2n cos (n − 2)x, etc. If n is even the middle term is constant and is the binomial coefficient alone, but if n is odd all terms are paired to form cosines.

Similar use of Euler's expression for the sine yields expressions for powers of sin x, distinct forms arising for even and odd powers.

For even powers of sin x we have

$$\sin^n x = \frac{(-1)^{\frac{n}{2}}}{2^{n-1}}\left[\cos nx - n\cos(n-2)x\right.$$

$$\left. + \frac{n(n-1)}{1.2}\cos(n-4)x + \ldots\ldots \pm R_n\right]. \qquad (9)$$

The expression in brackets is the same as in (8) when n is even except that the signs alternate. The sign before the bracket is positive if n is a multiple of 4, otherwise negative.

For odd powers of sin x we have

$$\sin^n x = \frac{(-1)^{\frac{n-1}{2}}}{2^{n-1}}\left[\sin nx - n\sin(n-2)x\right.$$

$$\left. + \frac{n(n-1)}{1\cdot 2}\sin(n-4)x + \ldots\right]. \qquad (10)$$

Here the number of terms is $\frac{n+1}{2}$, the last term involves sin x, the coefficients are all formed by the binomial law and the sign before the bracket is positive if n − 1 is a multiple of 4.

As an example consider the displacement s of the piston in a reciprocating engine or pump as a function of the crank angle θ. We have

$$s = L + r - r\cos\theta - \sqrt{L^2 - r^2\sin^2\theta},$$

where L is the length of connecting rod and r is radius of crank. We may write

$$\sqrt{L^2 - r^2 \sin^2\theta} = \frac{r}{p}(1 - p^2 \sin^2\theta)^{\frac{1}{2}}$$

where $pL = r$, and expanding by binomial theorem this becomes

$$\frac{r}{p}\left[1 - \frac{1}{2}p^2 \sin^2\theta - \frac{1}{8}p^4 \sin^4\theta - \frac{1}{16}p^6 \sin^6\theta + \ldots\right].$$

Applying equation (9) we have

$$\sin^2\theta = -\frac{1}{2}\cos 2\theta + \frac{1}{2},$$

$$\sin^4\theta = \frac{1}{8}\cos 4\theta - \frac{1}{2}\cos 2\theta + \frac{3}{8},$$

$$\sin^6\theta = -\frac{1}{32}\cos 6\theta + \frac{3}{16}\cos 4\theta - \frac{15}{32}\cos 2\theta + \frac{10}{32}.$$

Collecting results we obtain a Fourier series for s of the form

$$s = a_0 + a_1 \cos\theta + a_2 \cos 2\theta + a_4 \cos 4\theta + a_6 \cos 6\theta + \ldots$$

in which the constants are given up to the fifth power of p by the equations

$$a_0 = r\left(1 + \frac{p}{4} + \frac{3p^3}{64} + \frac{5p^5}{256}\right)$$

$$a_1 = -r$$

$$a_2 = -r\left(\frac{p}{4} + \frac{p^3}{16} + \frac{15p^5}{512}\right)$$

$$a_4 = r\left(\frac{p^3}{64} + \frac{3p^5}{256}\right)$$

$$a_6 = -r\left(\frac{p^5}{512}\right).$$

As another example consider the function

$$y = \frac{\sin x}{1 - p\cos x},$$

where p is small in comparison with unity. We may write

$$y = \sin x(1 + p\cos x + p^2\cos^2 x + p^3\cos^3 x + \ldots),$$

and from equation (8)

$$\cos^2 x = \frac{1}{2}\cos 2x + \frac{1}{2},$$

$$\cos^3 x = \frac{1}{4}\cos 3x + \frac{3}{4}\cos x,$$

and the function may be expressed as

$$y = \left(1 + \frac{p^2}{2}\right)\sin x + \left(p + \frac{3p^3}{4}\right)\sin x \cos x$$

$$+ \frac{p^2}{2}\sin x \cos 2x + \frac{p^3}{4}\sin x \cos 3x + \ldots .$$

Now the products here may be expressed as sums of harmonics as follows:

$$\sin x \cos x = \frac{1}{2}\sin 2x,$$

$$\sin x \cos 2x = \frac{1}{2}\sin 3x - \frac{1}{2}\sin x,$$

$$\sin x \cos 3x = \frac{1}{2}\sin 4x - \frac{1}{2}\sin 2x.$$

This gives for y the following expression, in which the coefficients are correct up to the third power of p.

$$y = \left(1 + \frac{p^2}{4}\right)\sin x + \left(p + \frac{p^3}{4}\right)\sin 2x$$

$$+ \frac{p^2}{4}\sin 3x + \frac{p^3}{8}\sin 4x + \ldots .$$

(c) If the methods of (a) and (b) above fail to yield results for a given function, still we may obtain approximately the values of the coefficients for the first few terms. The number of terms that may be obtained, and the reliability of the approximation, will depend upon the extent and character of the information at hand concerning the function.

Whether the function is given by an equation, by a graph or by

empirical data, the information at hand must yield a graph which represents the function from $x = 0$ to $x = 2\pi$ at least approximately. The area under the graph, proper attention being given to scales, will furnish a value for a_0 by equation (5). By the methods of Art. 4 the graphs of the functions in the integrals of (6) and (7) may be constructed for each value of m, and the areas under these curves will furnish the values of the integrals involved, and hence of the coefficients. These areas may be found by a planimeter, or by any other practical method, and the determination of any particular coefficient does not depend upon the determination of any other coefficient. Thus, for example, we may determine the seventh harmonic without finding any preceding harmonic.

The information at hand will yield also the values of the function, at least approximately, at regular intervals along the x-axis. These values are ordinates of the graph at the stated intervals. Various tabular methods are employed in different special applications of harmonic analysis, in which certain harmonics are computed by the use of ordinates at selected intervals. We illustrate here one of these methods, one of very general application, which is equivalent to the computation of the areas mentioned above by the trapezoidal rule

x°	y	cos x	y cos x	sin x	y sin x
0	9.3	1.000	9.30	0.000	0.00
30	15.0	0.866	12.99	0.500	7.50
60	17.4	0.500	8.70	0.866	15.07
90	23.0	0.000	0.00	1.000	23.00
120	37.0	−0.500	−18.50	0.866	32.04
150	31.0	−0.866	−26.85	0.500	15.50
180	15.3	−1.000	−15.30	0.000	0.00
210	4.0	−0.866	−3.46	−0.500	−2.00
240	−8.0	−0.500	4.00	−0.866	−6.93
270	13.0	0.000	0.00	−1.000	−13.20
300	−14.0	0.500	−7.10	−0.866	−12.30
330	−6.0	0.866	−5.20	−0.500	−3.00
sum	111.0		−41.36		55.68

x^o	y	cos 5x	y cos 5x	sin 5x	y sin 5x
0	9.3	1.000	9.30	0.000	0.0
30	15.0	−0.866	−12.99	0.500	7.50
60	17.4	0.500	8.70	−0.866	−15.07
90	23.0	0.000	0.00	1.000	23.00
120	37.0	−0.500	−18.50	−0.866	−32.04
150	31.0	0.866	26.85	0.500	15.50
180	15.3	−1.000	−15.30	0.000	0.00
210	4.0	0.866	3.46	−0.500	−2.00
240	−8.0	−0.500	4.00	0.866	−6.93
270	−13.0	0.000	0.00	−1.000	13.20
300	−14.0	0.500	−7.10	0.866	−12.30
330	−6.0	−0.866	5.20	−0.500	3.00
sum			3.62		−6.14

in Art. 45, thus approximating the integrals (5), (6) and (7). In one column we tabulate the values of x, usually in degrees, at which ordinates are to be taken; in the next column the corresponding values of $y = f(x)$; in other columns the values of sin x, cos x, sin 2x, cos 2x, etc., as far as it is desired to carry the work, and in further columns the products y sin x, y cos x, y sin 2x, y cos 2x, etc. The ordinate at $x = 0$ being included, that at $x = 2\pi$ is omitted provided that if these two are different, so that the periodic function is discontinuous, then the value at $x = 0$ is taken as the mean of the two values. Let M be the average of the values in any column, then $2\pi M$ is the value of the corresponding integral, and the corresponding constant may be found at once. The table given here contains the columns necessary to determine a_o, a_1, b_1, a_5, and b_5, illustrating the fact that the fifth harmonic may be found without preceding harmonics.

The sum of the 12 ordinates being 111.0, the average for the column is 9.25, therefore $a_o = 9.25$. Similarly the averages for y cos x and y sin x are −3.44 and 4.64, respectively, giving $a_1 = -6.88$ and $b_1 = 9.28$. Proceeding, we find $a_5 = 0.604$ and $b_5 = -1.02$.

In the use of this tabular method no attempt should be made to obtain a_m or b_m unless the number of ordinates used is greater than 2m. It may be noted that in the example given sin 6x is zero at

every value of x used, and cos 6x is alternately + 1 and −1 at the values used, so that the results for a_6 and b_6 would be quite unlike the results we should expect if more ordinates were employed, and are therefore unreliable. The more ordinates used the more reliably any given harmonic may be determined, but results are fairly reliable if the number of ordinates is at least 12 and exceeds 2m, as indicated above.

EXERCISES XLIX

1. Find the Fourier series which shall represent the function $y = 1$ from $x = 0$ to $x = \pi$ and $y = -1$ from $x = \pi$ to $x = 2\pi$. Note that the series is of form (4).

2. Integrate the series found in example 1 term by term, thus obtaining the series which represents $y = x$ from $x = 0$ to $x = \pi$, and $y = 2\pi - x$ from $x = \pi$ to $x = 2\pi$, except for the constant of integration. Express the constant of integration as a series of constants, by putting $y = 0$ when $x = 0$.

3. In the first illustration in (a) of Section 4 above we obtain a series for $y = x$ between $x = 0$ and $x = 1$. Show that differentiation is not valid for this series.

4. Integrate the series referred to in example 3 and determine the constant of integration as a series of constants by putting $x = 0$. Check by putting $x = 1$. Now by putting $x = \frac{1}{2}$ show that

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \ldots .$$

5. Given the function $y = (1 + p \sin x)^{\frac{3}{2}}$, where p is small compared to unity, find a Fourier series for the function, accurate up to the third power of p.

6. The polar equation of an ellipse is

$$\rho = \frac{2ep}{1 - e \cos \theta}$$

where e is less than unity. Express ρ as a Fourier series in θ, giving coefficients up to the 5th power of e.

7. In example 6 put $p = 5$ and $e = 0.1$, and compute values of ρ from the given equation for $\theta = 0, \frac{\pi}{6}, \frac{\pi}{3}$, etc. With these 12 values of ρ compute numerically a_0, a_2, a_3, and a_5, and compare with the values

found in 6 for the same values of p and e. Now compute a_0, a_2 and a_3 using only six values of ρ at 0, $\frac{\pi}{3}$, $\frac{2\pi}{3}$, etc., and compare with former results.

8. Analyze, as far as the given information permits, the periodic function six of whose ordinates are given by the table:

x.....	0°	60°	120°	180°	240°	300°
y.....	−0.85	0.95	0.72	2.75	−1.37	−2.20

9. Find the constant term and the first, third and fifth harmonics of the function given by the following data.

x.....	0°	30°	60°	90°	120°	150°
y.....	−18	−39	−39	−8	22	22
x.....	180°	120°	240°	270°	300°	330°
y.....	11	10	14	12	15	−1

10. Seven ordinates of a certain curve at intervals of $\frac{\pi}{6}$ are:

x.....	0°	30°	60°	90°	120°	150°	180°
y.....	0	676	660	940	1004	554	0

This information pertains to the first half of the cycle. Make three separate computations for the harmonics, based on three different assumptions:

(a) The series has no sine terms. Note that in this case the ordinates for the second half cycle are the given set repeated in reverse order, and that the average for any column of 12 values is the same as the average for the first six values in that column.

(b) The series has no cosine terms. Note that here the ordinates are repeated in reverse order but with opposite signs. Again the average for six values is the same as the average for 12 values in all but the y column.

(c) The series contains odd harmonics only. In this case the value at $x_0 + \pi$ is the negative of the value at x_0, so that the ordinates are repeated in the same order but with opposite signs. Here also the average of the six values is the same as for 12 values. Note that the computations in (a) and (b) are available here. There is no constant term.

11. A curve representing an alternating electric current has ordinates as follows:

x.....	0	10	20	30	40	50	60	70	80
y.....	5	4	20	22	19	25	29	29	30
x.....	90	100	110	120	130	140	150	160	170
y.....	38	46	38	41	50	32	30	33	7

Assuming that for the second half of the cycle these same values are repeated in the same order but with negative signs, find the first, 5th and 15th harmonics.

12. Find a Fourier series which shall represent the function $y = e^{-x}$ from $x = 0$ to $x = 2$, (a) by actual integration to determine coefficients; (b) using ordinates at intervals of 0.2. (Note that the ordinate at $x = 0$ must be taken as the mean of values at $x = 0$, and $x = 2$.)

BOOKS FOR SUPPLEMENTARY READING

Engineering Mathematics, C. P. Steinmetz (McGraw-Hill, 1911), Chapters I–V.

Hyperbolic Functions, James McMahon (Mathematical Monographs, Merriman and Woodward, No. 4, Wiley, 1906).

Advanced Calculus, W. F. Osgood (Macmillan, 1925), Chapters XVI and XX.

Advanced Calculus, E. B. Wilson (Ginn, 1912), Chapters VI, XVII and XVIII.

Mathematical Analysis, Goursat-Hedrick, Vol. I (Ginn, 1904), Chapter IX, also Vol. II, Part I (Ginn, 1916).

Integral Calculus, W. E. Byerly (Ginn, 1888), Chapters II and XVII.

Functions of a Complex Variable, T. S. Fiske (Mathematical Monographs, Merriman and Woodward, No. 11, Wiley, 1906).

Graphical and Mechanical Computation, Joseph Lipka (Wiley, 1918), Chapter VII.

The Calculus of Observations, Whitaker and Robinson (Blackie and Son, 1924), Chapters X and XIII.

Fourier's Series and Spherical Harmonics, W. E. Byerly (Ginn, 1893), especially Chapters II and III.

Fourier's Theorem and Harmonic Analysis, Albert Eagle (Longmans, 1925).

CHAPTER XII

FUNCTIONS OF SEVERAL VARIABLES

Art. 53. Geometrical Considerations

In previous chapters our attention has been given to functions of a single independent variable. In the present chapter we shall consider methods of treating functions which are best regarded as depending on two or more independent variables. If one or more equations, say m equations, express independent relations that hold among several variables, say n variables, the number of independent variables is $n - m$, for having assigned values to $n - m$ of the variables the m equations determine the values of the remaining m variables and all but one may be eliminated, leaving the one as a function of the $n - m$ variables selected as independent. For example the three equations

$$x = u - v, \quad y = u + v, \quad z = u^2 - v^2,$$

among the five variables involved, may be regarded as defining z as a function of the two independent variables x and y, for, given values of x and y, we may at once find values for u and v, and thus find a value for z. Also we may at once eliminate u and v and find $z = xy$, in which z is expressed directly in terms of the two independent variables.

In this we have regarded the several equations as simultaneous, but in many cases it is desirable to consider the relation expressed by each equation separately, later, perhaps, combining results. Thus, x, y and z, above may each be treated as functions of the two independent variables u and v.

The notation for functions of several variables is similar to that for a single variable, differing merely in that notations for the several variables are placed in evidence. Thus $z = f(x, y)$, $W = \phi(x, y, z)$ indicate that z is a function of x and y, and w of x, y and z.

Just as we have found the graphs of functions of a single variable, and the notions of plane analytic geometry very helpful in the study of those functions, so at this point we shall find it profitable to use the notions of three-dimensional geometry. The extension of the language of geometry to the case of four or more variables is sometimes desirable, and a geometry of four or more dimensions is entirely logical and useful, but, of course, not possible to realize visually.

1. **Coordinates and loci.** Given a relation among three variables, say $\phi(x, y, z) = 0$ or $z = f(x, y)$, we may interpret x, y and z as distances from three planes mutually at right angles, i.e., the three variables become the rectangular coordinates of a point in space. The origin O being chosen, three axes are determined through 0, xx′ horizontally, zz′ vertically and yy′ perpendicular to the plane of the first two (Fig. 103).

The point P as shown in the figure has its three coordinates positive. The seven other combinations of + and − signs with the three coordinates pertain to points placed in the other seven **octants** in sufficiently obvious manner. Any three real numbers, x, y, z, whether positive or negative, pertain to a definite point P = (x, y, z), and conversely any point P definitely determines its three distances, x from the yz-plane, y from the xz-plane, and z from the xy-plane.

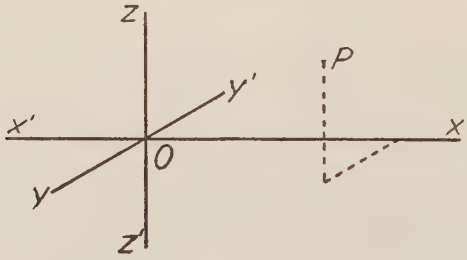


Fig. 103

If the coordinates, x, y and z, of a point P must fulfil a relation as $\phi(x, y, z) = 0$, the point P must have a definite z for each set of values assigned to x and y, and if z varies continuously with x and y the locus of P is a **surface.** For example, the **distance** from P to the origin O is $\sqrt{x^2 + y^2 + z^2}$ and if $x^2 + y^2 + z^2 = 16$, the point P remains at a distance 4 from O, and its locus is the surface of a sphere of radius 4, center at O.

Two surfaces, given by equations $\phi(x, y, z) = 0$, $f(x, y, z) = 0$, will in general intersect in a curve, the locus of points whose coordinates fulfill both equations. In general, then, two simultaneous equations in x, y and z determine a **space curve,** and conversely, any space curve may be determined by two equations, the equations of any two distinct surfaces that contain the curve. These two surfaces may be chosen in a multitude of ways, hence the equations of a space curve are not unique. Any equation of the form $\phi(x, y, z) + kf(x, y, z) = 0$ is satisfied by any x, y and z which satisfy $\phi(x, y, z) = 0$ and

$f(x, y, z) = 0$, and therefore contains the curve common to these surfaces. Very frequently k may be so chosen that the equation becomes simpler and easier of interpretation than either of the given equations. If one of the variables, say z, is eliminated, the resulting equation gives a relation between x and y which is true for all values of z, i.e., represents a cylindrical surface whose elements are parallel to the z-axis, and which contains the common curve.

As an example, consider the two surfaces

$$x^2 + y^2 + z^2 = 16, \quad x - z = 0.$$

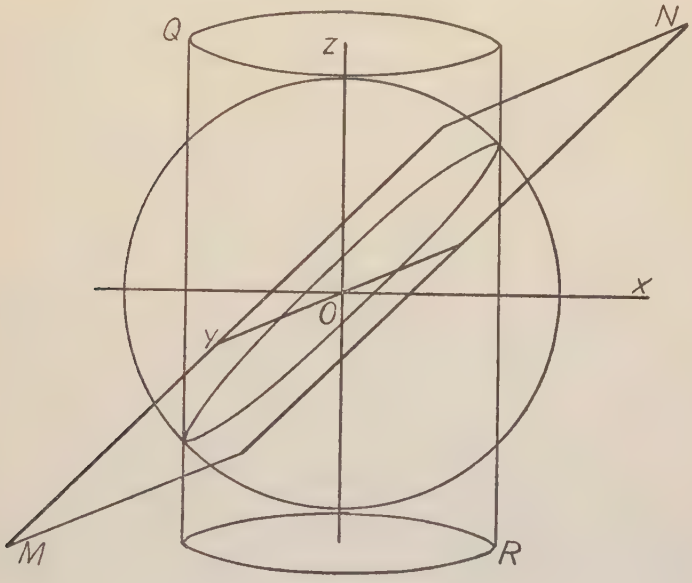


Fig. 104

The first is the sphere mentioned above, and the second is the plane bisecting the angle between the xy-plane and the yz-plane (Fig. 104). This plane, MN in the figure, cuts the sphere in a great circle. If we eliminate z we have $2x^2 + y^2 = 16$, an elliptical cylinder, QR in the figure, which also passes through the given great circle. If we merely add the given equations we have $x^2 + y^2 + z^2 + x - z = 16$, which is, in fact, the equation of another sphere of which the given circle is a small circle. Any sum or difference of any two of these four equations will give us another surface containing the given circle.

2. **Lines and planes.** If two points, (x_1, y_1, z_1) and (x_2, y_2, z_2), are

joined by a line, this line is a diagonal of a rectangular parallelopiped, the lengths of the edges of which are $x_2 - x_1$, $y_2 - y_1$, $z_2 - z_1$, the three edges being parallel respectively to the x-axis, the y-axis and the z-axis. The length of the diagonal is

$$l = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \tag{1}$$

The **direction cosines** of this diagonal are

$$\cos\alpha = \frac{x_2 - x_1}{l}, \quad \cos\beta = \frac{y_2 - y_1}{l}, \quad \cos\gamma = \frac{z_2 - z_1}{l}, \tag{2}$$

The angles α, β and γ are the angles that the diagonal makes with the three axes, or, what is the same thing by definition, with the three edges of the rectangular parallelopiped.

If any other point, (x, y, z), is on this same line, it must determine with (x_1, y_1, z_1), a rectangular parallelopiped of the same shape, so that

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}. \tag{3}$$

These equations then, are the equations of the line through the **two given points,** and are clearly equivalent to

$$\frac{x - x_1}{\cos\alpha} = \frac{y - y_1}{\cos\beta} = \frac{z - z_1}{\cos\gamma}, \tag{4}$$

the equations of a line determined by a **point and the direction** angles. If S is the distance from the point (x, y, z) to the fixed point (x_1, y_1, z_1), then the equations, taking account of (2) above, may be written in **parametric form,** S being the parameter,

$$x = x_1 + S\cos\alpha, \; y = y_1 + S\cos\beta, \; z = z_1 + S\cos\gamma \tag{5}$$

To determine the **angle between two lines,** we may, without loss of generality, suppose both lines to pass through O. Let OP (Fig. 105) have length l_1, direction angles α_1 β_1 γ_1, and let the coordinates of P be x_1, y_1, z_1. Let OQ have length l_2 direction angles α_2, β_2, γ_2, and let PQ be perpendicular to OQ. Now if θ is the angle between the lines

$$\cos\theta = \frac{OQ}{OP} = \frac{l_2}{l_1}.$$

The coordinates, x_1, y_1, z_1, form a series of line segments joining O to P, and the sum of the projections of these segments on OQ is the same as the projection of OP and OQ, namely l_2. We have then,

$$x_1 \cos \alpha_2 + y_1 \cos \beta_2 + z_1 \cos \gamma_2 = l_2.$$

But $x_1 = l_1 \cos \alpha_1$, $y_1 = l_1 \cos \beta_1$, $z_1 = l_1 \cos \gamma_1$, and substituting these values and dividing by l_1 we have

$$\cos \alpha_1 \cos \alpha_2 + \cos \beta_1 \cos \beta_2 + \cos \gamma_1 \cos \gamma_2 = \frac{l_2}{l_1} = \cos \theta. \quad (6)$$

In particular, if the two lines are **parallel**, $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$, $\gamma_1 = \gamma_2$, and the relation becomes

$$\cos^2 \alpha_1 + \cos^2 \beta_1 + \cos^2 \gamma_1 = \cos 0 = 1, \quad (7)$$

a relation that is identically fulfilled by the directions cosines of any line.

If the two lines are **perpendicular** to each other, we must have $\cos \theta = 0$, or

$$\cos \alpha_1 \cos \alpha_2 + \cos \beta_1 \cos \beta_2 + \cos \gamma_1 \cos \gamma_2 = 0. \quad (8)$$

If a point (x, y, z) is always in such position that its projection on a line OP is at P, the locus of the point is a plane perpendicular to OP at P, and if the length OP is p, the equation expressing the condition is

$$x \cos \alpha + y \cos \beta + z \cos \gamma = p, \quad (9)$$

where α, β and γ are the directions angles of the line OP. This equation is known as the **normal equation** of the plane.

Any linear equation in x, y and z, as $Ax + By + Cz + D = 0$ represents a plane, for there is always a rectangular parallelopiped whose edges are A, B and C, the diagonal of which has direction cosines $\frac{A}{R'}, \frac{B}{R'}, \frac{C}{R'}$, where $R = \sqrt{A^2 + B^2 + C^2}$, and dividing the equation by R and transposing the constant term it becomes the normal equation of a plane. In particular, the equation

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1 \quad (10)$$

represents a plane whose **intercept** on the x-axis is a, on the y-axis is b, and on the z-axis is c. Also the equation

$$(x - x_1)\cos\alpha + (y - y_1)\cos\beta + (z - z_1)\cos\gamma = 0 \qquad (11)$$

represents a plane **through the point** (x_1, y_1, z_1) **and perpendicular to a line** whose direction angles are α, β and γ.

The equation

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0, \qquad (11a)$$

represents a plane **through the three points** (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3), for it is a linear equation in x, y and z and is obviously

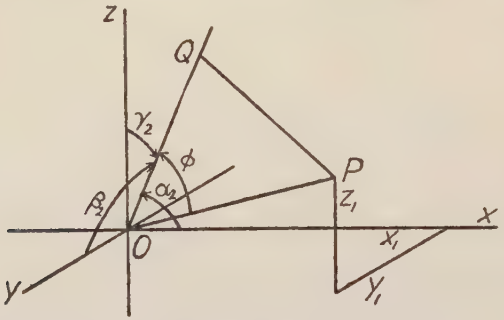


Fig. 105

satisfied if the coordinates of one of the three points are substituted for the variables. Noticing the analogy between this equation and (6′) of Art. 31, one may seek a formula for volume of a tetrahedron analogous to (3′) of Art. 30. Without taking space for the proof, we may state that one-sixth of the determinant written above is the **volume of the tetrahedron** whose vertices are the four points whose coordinates appear. The value will, of course, be positive or negative according to the order of the four points.

Since two planes determine a line, any two linear equations represent a line, and furnish relations which easily lead to equations of form (3), (4) or (5).

For example, given the equations

$$3x - 5y + z = 3,$$

$$x + 7y - z = 1,$$

we may add, obtaining,

$$4x + 2y = 4, \text{ or } y = -2x + 2.$$

Now eliminating x, we have

$$26y = 4z, \text{ or } y = \frac{2z}{13}.$$

We may write then,

$$-2x + 2 = \frac{y}{1} = \frac{2z}{13},$$

or

$$\frac{x - 1}{1} = \frac{y}{-2} = \frac{z}{-13},$$

and this form may be reduced to form (4) by dividing the three denominators by the square root of the sum of their squares. This last step, however, is awkward, and usually not desirable. To use form (3) it is only necessary to find two sets of values of x, y and z which satisfy the two given equations, and use these values as x_1, y_1, z_1, and x_2, y_2, z_2.

The **angle between two planes**

$$A_1x + B_1y + C_1z + D_1 = 0,$$

$$A_2x + B_2y + C_2z + D_2 = 0.$$

is the same as that between the normals to these planes, and therefore denoting the angle by θ,

$$\cos\theta = \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{(A_1^2 + B_1^2 + C_1^2)(A_2^2 + B_2^2 + C_2^2)}}. \tag{12}$$

If the planes are **perpendicular** to each other we must have

$$A_1A_2 + B_1B_2 + C_1C_2 = 0. \tag{13}$$

Comparing equation (11) with (9) we see that the perpendicular distance from the origin to a plane through (x_1, y_1, z_1) is

$$x_1 \cos\alpha + y_1 \cos\beta + z_1 \cos\gamma,$$

where α, β and γ are the direction angles of the normal to the plane. The equation of any plane parallel to this is

$$x \cos\alpha + y \cos\beta + z \cos\gamma = p,$$

where p is its distance from the origin. The **perpendicular distance from** (x_1, y_1, z_1) **to this latter plane** is therefore,

$$d = x_1 \cos \alpha + y_1 \cos \beta + z_1 \cos \gamma - p. \tag{14}$$

To illustrate, if we wish to find the distance from $(2, 5, -3)$ to the plane $3x - 4y + 12z = 20$, we have, since $3^2 + 4^2 + 12^2 = 13^2$,

$$d = \frac{3(2) - 4(5) + 12(-3) - 20}{13} = -\frac{70}{13}.$$

3. **Loci of higher degree than the first.** The study of surfaces whose equations are of the second or higher degree is facilitated by considerations of **symmetry** with respect to the coordinate axes and planes, also by consideration of the curves in which the surface is cut by certain **systems of planes.**

A locus is symmetrical with respect to the origin if all the variables may have their signs changed without effect on the equation. It is symmetrical with respect to the axis of one variable if the other two variables may be so changed, and it is symmetrical to the plane of two variables if the remaining variable may be so changed. For example the surface $x^2 + 3y^2 + 4yz = 10$ is symmetrical with respect to the origin, with respect to the x-axis, and with respect to the yz-plane, but not with respect to the y-axis, the z-axis, the xz-plane nor the xy-plane.

The systems of planes whose intersections with a surface are most convenient and helpful are those parallel to the coordinate planes, with equations of the form $x = k$, $y = k$, $z = k$, and in particular, setting $x = 0$, $y = 0$, and $z = 0$ in the equation gives the intersections with the coordinate planes, called the **traces** of the surface. For instance the trace of the surface in the example above on the yz-plane is the hyperbola $3y^2 + 4yz = 10$, on the xz-plane is the pair of line $x^2 = 10$, and on the xy-plane is the ellipse $x^2 + 3y^2 = 10$. To further study this surface we may consider sections by planes $z = k$, giving the system of ellipses in parallel planes

$$x^2 + 3y^2 + 4ky = 10,$$

and we may proceed in like manner to sections by planes $x = k$, and $y = k$.

Any equation of the form

$$x^2 + y^2 = f(z) \tag{15}$$

will have for sections with $z = k$ circles with centers on the z-axis. Such a surface is called a **surface of revolution about the z-axis.** Planes

through the z-axis will cut this surface in **meridians,** whose form is that of the trace in the xz-plane, $x^2 = f(z)$. Of course we may make a similar discussion for the other axes by interchanging variables.

Surfaces of the second degree are called **quadric surfaces.** With suitable choice of axes they reduce to three types. First the **cones,**

$$\frac{x^2}{a^2} \pm \frac{y^2}{b^2} \pm \frac{z^2}{c^2} = 0. \tag{16}$$

with vertex at origin, with axis along the axis of that variable whose square differs in sign from the other two. If the three squares are of like signs, the locus has a single real point, the origin. Second, the **ellipsoids** and **hyperboloids,**

$$\pm\frac{x^2}{a^2} \pm \frac{y^2}{b^2} \pm \frac{z^2}{c^2} = 1, \tag{17}$$

ellipsoid if all terms are positive, **hyperboloid of one sheet** if one term is negative, **hyperboloid of two sheets** if two terms are negative, imaginary if all three signs are negative. Third, the **paraboloids**

$$\pm\frac{x^2}{a^2} \pm \frac{y^2}{b^2} = cz, \tag{18}$$

elliptic if the ambiguous signs agree, **hyperbolic** if they differ.

The names of these types of surfaces are easily justified by application of the methods of study indicated above.

A very useful method of representing a surface is the **parametric** form

$$x = f_1(u, v), \quad y = f_2(u, v), \quad z = f_3(u, v), \tag{19}$$

from which three equations a single relation among x, y and z might be found by eliminating the two parameters u and v. Similarly the equations

$$x = f_1(t), \quad y = f_2(t), \quad z = f_3(t), \tag{20}$$

determine a curve in space, two relations among x, y and z resulting from elimination of t from first one pair then another. For example, a sphere is given by

$$x = r \cos u \cos v,$$

$$y = r \cos u \sin v,$$

$$z = r \sin u,$$

where u may be taken as latitude and v as longitude. If we assign a relation to u and v, these equations give a curve on the sphere. Meridians are given by $v = k$, while parallels are given by $u = k$, and a spiral cutting all meridians at a fixed angle is given by $u = mv$. Any **surface of revolution** about the z-axis may be represented in the form

$$x = u \cos v, \quad y = u \sin v, \quad z = f(u) \tag{21}$$

while a **helicoidal** or screw surface is given by

$$x = u \cos v, \quad y = u \sin v, \quad z = f(u) + mv. \tag{22}$$

If the arbitrary function f(u) is always zero, this becomes the ordinary **helicoid** which may be generated by a straight line perpendicular to the z-axis. If now we put $u = r$, a constant, these equations become

$$x = r \cos v, \quad y = r \sin v, \quad z = mv, \tag{23}$$

the equations of a **helix,** the form of a screw thread on a cylinder of radius r and having pitch $2\pi m$.

EXERCISES L

1. Find the distance from (3, 1, 1) to (−2, −1, 0).

2. Find the equations of a cylinder through the curve of intersection of $x^2 - 3y^2 + 2z^2 - z = 0$ and $2x^2 + y^2 + z = 0$ with generators parallel to (a) the x-axis; (b) the y-axis; (c) the z-axis.

3. Find the direction cosines of the line joining (2, 3, 3) and (1, 1, 1) if the line is directed upward.

4. Find the equations of the line joining (2, 3, 3) and (1, 1, 1).

5. Find the coordinate of a point on above line 3 units from (1, 1, 1) in the positive direction.

6. Find equations of the line through (5, −2, 4) and (3, 1, 2) in each of the forms (3), (4) and (5) given above.

7. Find the equation of a plane through the point (6, 2, 3) and perpendicular to the line joining this point to the origin.

8. Find the direction cosines of a normal to the plane, $5x - 2y - 2z + 3 = 0$.

9. Find in form (4) equations of the intersection of $3x + y - 7z + 11 = 0$ and $2x + 4y + z - 3 = 0$.

10. Find in parametric form equations of the intersection of $2y - 3z + 5 = 0$ and $7x - 4y - z - 10 = 0$.

11. Find the angle at which the line in exercise 10 and the intersection of $y + 1 = 0$ and $2x + 3y + 2z - 1 = 0$ intersect.

12. Find the equation of the plane through (1, 4, 1) perpendicular to the line $x + 3y + 5z + 6 = 0$, $y + 2x - 1 = 0$.

13. Write the equation of a plane through the points (2, 1, 7), (1, −2, 7) and (4, 3, −1).

14. The tetrahedron whose vertices are (0, 0, 0), (1, 0, 0), (0, 1, 0) and (0, 0, 1) obviously has a volume $\frac{1}{6}$. Find this as one-sixth of a determinant of the form of (11a).

15. Find the volume of the tetrahedron whose vertices are (1, 3, 2), (2, 7, 5), (−3, 4, 6) and (3, 0, 2).

16. Show that the four points (5, 6, 7), (2, 7, 4), (3, 4, −3) and (1, 6, −1) lie in one plane.

17. Find the angle between the line $x - 2y - 8 = 0$, $3y + z + 8 = 0$ and the plane $3x + y - 2z + 7 = 0$.

18. Find the distance from point (4, 3, 7) to the plane $5x - 2y + 4z - 7 = 0$.

19. Find the equation of the plane passing through the line $x - y + z = 0$, $2x + y + 3z = 0$ and also through the point (1, 2, −1).

20. Find the locus of a point such that the sum of its distances from (5, 0, 0) and (−5, 0, 0) is 15. Discuss the surface.

21. Find the locus of a point equidistant from a fixed plane and a fixed point not in the fixed plane.

22. Find the equation of the surface formed by revolving the plane curve $x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}$ about the y-axis.

23. Show that $x = t^2$, $y = 2t$, $z = t$ is a plane section of a parabolic cylinder.

24. Show that the surface $x^2 = uv^2$, $y = u(1 - v^2)$, $z^2 = 4u$ is a cone of revolution about the z-axis.

25. Determine the nature of the curve $x = t \cos t$, $y = t \sin t$, $z = t$.

Art. 54. Partial and Total Derivatives

1. **Partial derivatives.** First let us consider a function of two independent variables, $z = f(x, y)$. Let us fix attention on the case when y is regarded as a constant and consider z as a function of x. Let y have the constant value y_0, and consider the derivative of z with respect to x at $x = x_0$. The function notation displays both variables, and to indicate that the derivative is taken regarding x as the only independent variable we use a distinctive notation for the derivative, $\frac{\partial z}{\partial x}$, and call it the **partial derivative** of z with respect to x.

We have

$$\frac{\Delta z}{\Delta x} = \frac{f(x_o + \Delta x, y_o) - f(x_o, y_o)}{\Delta x} \tag{1}$$

and the limit of this fraction, as Δx approaches zero, is the value of $\frac{\partial z}{\partial x}$ at (x_o, y_o). Of course x_o and y_o may have any values on the respective ranges, so that the partial derivative is a function of x_o and y_o, and may be found without specifying their values. The subscripts may be omitted, leaving the partial derivative a function of x and y.

Similarly we have for $\frac{\partial z}{\partial y}$ the limit of

$$\frac{\Delta z}{\Delta y} = \frac{f(x_o, y_o + \Delta y) - f(x_o, y_o)}{\Delta y} \tag{2}$$

and this also is a function of x_o and y_o, or of x and y, since x_o and y_o are any values of x and y.

As an illustration take $z = x^2 - 3xy + y^3$. We have $\frac{\partial z}{\partial x} = 2x - 3y$, $\frac{\partial z}{\partial y} = 3x + 3y^2$. These are found by the rules for finding derivatives developed in Chapter III, which are applicable here if one letter is considered as a constant.

Again, if w is a function of three variables,

$$w = f(x, y, z),$$

then the partial derivative of w with respect to x is the derivative found by regarding y and z as constant, treating w as a function of the single independent variable x. This partial derivative, $\frac{\partial w}{\partial x}$, is a function of the three variables x, y and z. Similar significance is attached to the notations $\frac{\partial w}{\partial y}$ and $\frac{\partial w}{\partial z}$. The ideas are easily extended to functions of any number of independent variables.

2. **Total derivatives and differentials.** Consider again a function of two variables,

$$z = f(x, y)$$

and suppose an increment Δz in z to be due to simultaneous incre-

ments Δx and Δy in the independent variables. We have then, supposing x_o, y_o, and z_o to be the initial values of the variables,

$$z = f(x_o + \Delta x, y_o + \Delta y) - f(x_o, y_o),$$

and adding and subtracting the quantity $f(x_o, y_o + \Delta y)$, and dividing by Δx, we obtain

$$\frac{\Delta z}{\Delta x} = \frac{f(x_o + \Delta x, y_o + \Delta y) - f(x_o, y_o + \Delta y)}{\Delta x}$$

$$+ \frac{f(x_o, y_o + \Delta y) - f(x_o, y_o)}{\Delta y} \cdot \frac{\Delta y}{\Delta x}.$$

The first expression on the right is the same as the right hand of (1) in 1 above, except that the function values are taken at $y_o + \Delta y$ instead of at y_o, and as Δx approaches zero, $y_o + \Delta y$ approaches y_o, so that the limit of this function is the partial derivative of z with respect to x at (x_o, y_o). The second term, comparing with (2) in 1, clearly has for limit the value at (x_o, y_o) of the quantity $\frac{\partial z}{\partial y}\frac{dy}{dx}$, where $\frac{dy}{dx}$ is the derivative of y with respect to x, in the ordinary sense, based on their mode of variation as the point (x, y) moves from $(x_o y_o)$. Now denoting the limit of $\Delta z/\Delta x$ by dz/dx, we have

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dx}, \tag{3}$$

and this quantity is called the **total derivative** of z with respect to x. It should be noted that the Δz that enters here is due to both increments Δx and Δy, while the Δz that enters in the definition of the partial derivatives is due to only one of these increments. The total derivative dz/dx depends for its value not only on x and y but also on dy/dx. For example, if as before

$$z = x^2 - 3xy + y^3,$$

then

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx} = 2x - 3y + (3y^2 - 3x)\frac{dy}{dx}.$$

The same result may be obtained by taking the derivative of the right hand side as in the case of implicit functions, Art. 13.

Reasoning by analogy, we have

$$\frac{dz}{dy} = \frac{\partial z}{\partial y} + \frac{\partial z}{\partial x}\frac{dx}{dy}, \tag{4}$$

and the notion of total derivative is easily extended to functions of any number of variables. If we take

$$w = f(x, y, z, u)$$

then

$$\frac{dw}{dx} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial y}\frac{dy}{dx} + \frac{\partial w}{\partial z}\frac{dz}{dx} + \frac{\partial w}{\partial u}\frac{du}{dx}, \tag{5}$$

and similar expressions are obtained for $\frac{dw}{dy}, \frac{dw}{dz}, \frac{dw}{du}$.

If we use differential notation, and remember that the ratio of differentials is the derivative, we may write, by (3) or (4) above

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy. \tag{6}$$

This expression is known as the **total differential** of z. From (5) above we have for the total differential of w

$$dw = \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz + \frac{\partial w}{\partial u}du. \tag{7}$$

3. **Geometrical interpretation.** In the case of a function of two independent variables the partial and total derivatives and total differentials are very simply interpreted in geometry. Let the equation $z = f(x, y)$ represent a surface, a portion of which is bounded by the irregular curve (Fig. 106). The plane $x = x_0$ cuts this surface in the curve PB, whose slope at P is

$$\lim_{\Delta y \to 0}\frac{\Delta z}{\Delta y} = \lim_{PX \to 0}\frac{YB}{PY} = \left[\frac{\partial z}{\partial y}\right]_P = \frac{\partial f(x_0, y_0)}{\partial y_0} = \frac{YN}{PY}$$

where PY is understood parallel to the y-axis and PN tangent to the curve PB at P.

Similarly $y = y_0$ cuts the surface in the curve PA, whose slope at P is

$$\lim_{\Delta x \to 0}\frac{\Delta z}{\Delta x} = \lim_{PX \to 0}\frac{XA}{PX} = \left[\frac{\partial z}{\partial x}\right]_P = \frac{\partial f(x_0, y_0)}{\partial x_0} = \frac{XM}{PX}.$$

Now for the total differential, if $dx = PX$ and $dy = PY$ we have $XM = \frac{\partial z}{\partial x}\,dx$, $YN = \frac{\partial z}{\partial y}\,dy$, and, PXQY being a rectangle in a horizontal plane and PMTN being a parallelogram in the plane tangent to the surface at P, it is clear geometrically $XM + YN = QT$, hence

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy = QT.$$

Again the Δz due to $\Delta x = PX$ and $\Delta y = PY$ is QC. If the point $C = (x_o + \Delta x,\ y_o + \Delta y,\ z_o + \Delta z)$ is understood to approach $P =$

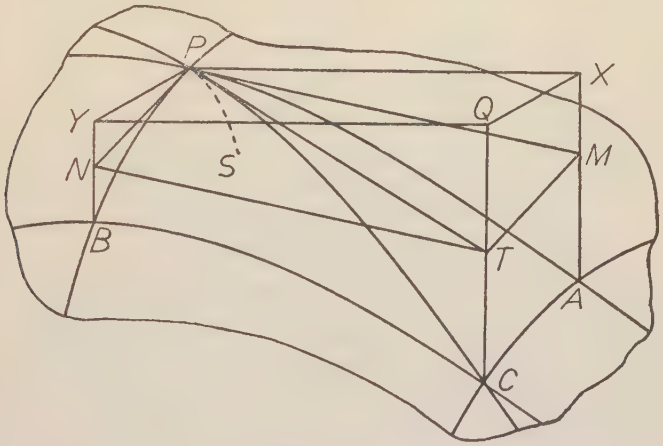


Fig. 106

$(x_o,\ y_o,\ z_o)$ along the curve PC, cut from the surface by the plane PQC, then since PT is tangent to this curve

$$\frac{dz}{dx} = \lim_{\Delta x \to o} \frac{\Delta z}{\Delta x} = \lim_{PX \to o} \frac{QC}{PX} = \frac{QT}{PX}.$$

But geometrically we see that

$$\frac{QT}{PX} = \frac{XM}{PX} + \frac{YN}{PX} = \frac{XM}{PX} + \frac{YN}{PY} \cdot \frac{PY}{PX} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dx}.$$

We have specified the path PC to simplify the statements above. It should be clear that if the point moves on any other curve, PS, of the surface, having PT as tangent at P, the ratio limits will be as above. The total derivative at a point P, then, depends not on the

curve PS, or PC, but on the direction of the curve at P. The partial derivatives are related to the tangent plane PMTN, while the total derivative is related also to the direction of the line PT in this plane.

4. **Partial derivatives of higher order.** As has been stated, the partial derivatives of a function of several variables are themselves functions of these variables, and these functions have partial and total derivatives. The successive partial derivatives of z with respect to x are:

$$\frac{\partial z}{\partial x'} \quad \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^{2'}} \quad \frac{\partial}{\partial x}\left(\frac{\partial^2 z}{\partial x^2}\right) = \frac{\partial^3 z}{\partial x^{3'}} \text{ etc.,}$$

and similarly with respect to any other independent variable. Again we have

$$\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial y \partial x'} \quad \frac{\partial}{\partial y}\left(\frac{\partial^2 z}{\partial x^2}\right) = \frac{\partial^3 z}{\partial y \partial x^2}.$$

If w is a function of x, y and z, we may have, among others

$$\frac{\partial^2 w}{\partial x\ \partial y'} \quad \frac{\partial^3 w^2}{\partial x\ \partial y\ \partial z'} \quad \frac{\partial^3 w}{\partial z\ \partial x\ \partial y'} \quad \frac{\partial^4 w}{\partial z\ \partial y^2\ \partial x'}$$

each with its obvious interpretation, the last for example, being found by taking the partial derivative of w with respect to x, the second partial of this with respect to y, and the partial of this with respect to z.

5. **Functions given indirectly.** Suppose relations given among several variables, as

$$f(x, y, z, w) = 0,$$

$$\phi(x, y, z, w) = 0. \tag{8}$$

If these equations are taken as simultaneous, the number of independent variables is equal to the total number of variables less the number of equations. Usually the conditions of the problem indicate a choice of independent variables, and, they being chosen, we may solve for the dependent variables, obtaining explicit functions as treated above, or we may treat the equations as defining the functions implicitly. Regarding x and y as independent variables, and z and w as functions of these, we may differentiate partially with respect to x.

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x} = 0,$$

$$\frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial z}\frac{\partial z}{\partial x} + \frac{\partial \phi}{\partial w}\frac{\partial w}{\partial x} = 0, \tag{9}$$

in which the partial derivatives of the functions f(x, y, z, w) and $\phi(x, y, z, w)$ are derived as functions of the four variables, while the partial derivatives of z and w can only be indicated by the symbols. The two equations may now be solved for these latter quantities, giving

$$\frac{\partial z}{\partial x} = \frac{\begin{vmatrix} \frac{\partial f}{\partial w} & \frac{\partial f}{\partial x} \\ \frac{\partial \phi}{\partial w} & \frac{\partial \phi}{\partial x} \end{vmatrix}}{\begin{vmatrix} \frac{\partial f}{\partial z} & \frac{\partial f}{\partial w} \\ \frac{\partial \phi}{\partial z} & \frac{\partial \phi}{\partial w} \end{vmatrix}}, \quad \frac{\partial w}{\partial x} = \frac{\begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial z} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial z} \end{vmatrix}}{\begin{vmatrix} \frac{\partial f}{\partial z} & \frac{\partial f}{\partial w} \\ \frac{\partial \phi}{\partial z} & \frac{\partial \phi}{\partial w} \end{vmatrix}}.$$

Similarly we may find the partial derivatives of z and w with respect to y. To find the total derivative of w with respect to x we have

$$\frac{dw}{dx} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial y}\frac{dy}{dx}. \tag{10}$$

We may find the necessary partial derivatives as just indicated or we may proceed at once to differentiate equations (8).

$$\begin{aligned} \frac{df}{dx} &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dy} + \frac{\partial f}{\partial w}\frac{dw}{dx} = 0, \\ \frac{d\phi}{dx} &= \frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y}\frac{dy}{dx} + \frac{\partial \phi}{\partial z}\frac{dz}{dx} + \frac{\partial f}{\partial w}\frac{dw}{dx} = 0. \end{aligned} \tag{11}$$

Eliminating $\frac{dz}{dx}$ and solving for $\frac{dw}{dx}$ the desired relation (10) will be found. The derivative $\frac{dy}{dx}$ must remain arbitrary, unless some relation is given from which it may be computed, but in this case we have essentially only one independent variable.

Of course the above discussion applies to a single function as well as to several functions, and indicates also the method of treating parametric equations.

As an example, let

$$xy + 3x^2z - yz^3 = 0 \tag{12}$$

define z as a function of x and y, and let it be required to find the functions

$$\frac{\partial z}{\partial x}, \frac{dz}{dy}, \frac{\partial^2 z}{\partial x\, \partial y}.$$

First we have

$$y + 6xz + 3x^2 \frac{\partial z}{\partial x} - 3yz^2 \frac{\partial z}{\partial x} = 0 \tag{13}$$

from which

$$\frac{\partial z}{\partial x} = \frac{y + 6xz}{3yz^2 - 3x^2}. \tag{14}$$

Again, taking total derivatives term by term,

$$x + y\frac{dx}{dy} + 3x^2 \frac{dz}{dy} + 6xz\frac{dx}{dy} - 3yz^2 \frac{dz}{dy} - z^3 = 0, \tag{15}$$

which gives

$$\frac{dz}{dy} = \frac{x - z^3}{3yz^2 - 3x^2} + \frac{y + 6xz}{3yz^2 - 3x^2}\frac{dx}{dy}. \tag{16}$$

Now differentiating (13) partially with respect to y,

$$1 + 6x\frac{\partial z}{\partial y} + 3x^2 \frac{\partial^2 z}{\partial y\, \partial x} - 3yz^2 \frac{\partial^2 z}{\partial y\, \partial x} - 3z^2 \frac{\partial z}{\partial x}$$

$$- 6yz \frac{\partial z}{\partial x}\frac{\partial z}{\partial y} = 0,$$

and substituting the values of $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$, the latter being the first fraction on the right in (16), we may solve, for $\frac{\partial^2 z}{\partial y\, \partial x}$.

As another example, consider the helicoid $x = u \cos v$, $y = u \sin v$, $z = mv$.

Let us regard these equations as defining z as a function of x and y, u and v being auxiliary variables. To find $\frac{\partial z}{\partial x}$ we take partial derivatives with respect to x, considering z, u and v as functions of x and y.

$$1 = \frac{\partial u}{\partial x}\cos v - u \sin v \frac{\partial v}{\partial x},$$

$$0 = \frac{\partial u}{\partial x}\sin v + u \cos v \frac{\partial v}{\partial x}, \quad \frac{\partial z}{\partial x} = m \frac{\partial v}{\partial x},$$

whence

$$\frac{\partial z}{\partial x} = -m \frac{\sin v}{u}.$$

EXERCISES LI

1. Given the function,

$$z = x^3 - xy^2 + 2y,$$

find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ at (2, 0). Also find $\frac{dz}{dx}$ at (2, 0), assuming that x and y are related by the equation $y = x^2 - 4$. Obtain this last result again by substituting $x^2 - 4$ for y in the expression for z, thus making it a function of x only.

2. If $z = \sin(x + y)$, show that $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y}$, and that

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial^2 z}{\partial x\, \partial y} = \frac{\partial^2 z}{\partial y\, \partial x} = \frac{\partial^2 z}{\partial y^2}$$

3. If $pv = kz$, find $\frac{\partial z}{\partial p}, \frac{\partial z}{\partial v}$ and the total differential of z, considering k as a constant.

4. Show that if y is defined as a function of x by an implicit equation, $f(x, y) = 0$, then

$$\frac{dy}{dx} = -\frac{\frac{\partial u}{\partial x}}{\frac{\partial u}{\partial y}}$$

where $u = f(x, y)$.

5. If $w = \frac{x}{y} + \frac{y}{z} - xz$, find $\frac{\partial^2 w}{\partial x\, \partial y}, \frac{\partial^2 w}{\partial z^2}$, and $\frac{\partial^2 w}{\partial z\, \partial x}$.

6. If $w = z^{xy}$ find the total differential of w. (Take logarithms of both sides before taking differentials).

7. Let z be determined as a function of x and y by the implicit equation

$$xy^2 + \sin^{-1}\frac{y}{z} = 0,$$

and find $\frac{\partial z}{\partial y}$, $\frac{\partial z}{\partial x}$ and dz.

8. Find $\frac{dz}{dt}$ if $z = x^2 + y^2$, $x = t$, $y = t^3$.

9. If $z = f\left(\frac{x}{y}\right)$, prove $x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = 0$.

10. If $\sin(x + y) - e^{xy} = 0$, find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

11. Given $u = \rho^2 e^{2\theta}$, $x = \rho\cos\theta$, $y = \rho\sin\theta$, find $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, regarding ρ and θ as auxiliary variables.

Art. 55. Applications of Partial and Total Derivatives

1. **Tangents and normals.** Let a surface be defined by the equation

$$f(x, y, z) = 0, \tag{1}$$

and let (x_o, y_o, z_o) be a point on this surface, then the equation

$$(x - x_o)\left(\frac{\partial f}{\partial x}\right)_o + (y - y_o)\left(\frac{\partial f}{\partial y}\right)_o + (z - z_o)\left(\frac{\partial f}{\partial z}\right)_o = 0 \tag{2}$$

is the equation of the **tangent plane** at the point. For, being of the first degree in x, y and z it is a plane; it must go through the point in question; and, since it gives the same partial derivatives of z with respect to x and y as does the given equation at the point, it must contain lines tangent to the intersections of (1) with $x = x_o$ and $y = y_o$ at this point.

If the equation is explicit, as

$$z = f(x, y) \tag{3}$$

the tangent plane becomes

$$z - z_o = \left(\frac{\partial z}{\partial x}\right)_o (x - x_o) + \left(\frac{\partial z}{\partial y}\right)_o (y - y_o), \tag{4}$$

while the parametric equations for a surface,

$$x = f_1(u, v), \quad y = f_2(u, v), \quad z = f_3(u, v) \tag{5}$$

yield the equations for the tangent plane

$$\begin{aligned} x - x_0 &= \left(\frac{\partial x}{\partial u}\right)_0 (u - u_0) + \left(\frac{\partial x}{\partial v}\right)_0 (v - v_0), \\ y - y_0 &= \left(\frac{\partial y}{\partial u}\right)_0 (u - u_0) + \left(\frac{\partial y}{\partial v}\right)_0 (v - v_0), \\ z - z_0 &= \left(\frac{\partial z}{\partial u}\right)_0 (u - u_0) + \left(\frac{\partial z}{\partial v}\right)_0 (v - v_0). \end{aligned} \tag{6}$$

If we eliminate $u - u_0$ and $v - v_0$ from these equations we have the equation of the tangent plane in the form

$$(x - x_0)\left|\begin{matrix} \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{matrix}\right|_0 + (y - y_0)\left|\begin{matrix} \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \end{matrix}\right|_0 + (z - z_0)\left|\begin{matrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{matrix}\right|_0 = 0, \tag{7}$$

where the zero subscript indicates that the determinant expression is evaluated at (x_0, y_0, z_0).

A **line normal to a surface** at a point (x_0, y_0, z_0) is perpendicular to the tangent plane at this point, and may be given in the forms

$$\frac{x - x_0}{\left(\frac{\partial f}{\partial x}\right)_0} = \frac{y - y_0}{\left(\frac{\partial f}{\partial y}\right)_0} = \frac{z - z_0}{\left(\frac{\partial f}{\partial z}\right)_0}, \tag{8}$$

$$\frac{x - x_0}{\left(\frac{\partial z}{\partial x}\right)_0} = \frac{y - y_0}{\left(\frac{\partial z}{\partial y}\right)_0} = \frac{z - z_0}{-1}, \tag{9}$$

$$\frac{x - x_0}{A} = \frac{y = y_0}{B} = \frac{z - z_0}{C}, \tag{10}$$

where A, B and C are respectively the first, second and third determinants in (7), the three forms, (8), (9) and (10) pertaining respectively to the forms (1), (3) and (5).

If a curve in space is given as the intersection of two surfaces,

$$f(x, y, z) = 0, \quad \phi(x, y, z) = 0, \tag{11}$$

then the line tangent to the curve at (x_o, y_o, z_o) is the intersection of the tangent planes to the surfaces at this point,

$$(x - x_o)\left(\frac{\partial f}{\partial x}\right)_o + (y - y_o)\left(\frac{\partial f}{\partial y}\right)_o + (z - z_o)\left(\frac{\partial f}{\partial z}\right)_o = 0,$$

$$(x - x_o)\left(\frac{\partial \phi}{\partial x}\right)_o + (y - y_o)\left(\frac{\partial \phi}{\partial y}\right)_o + (z - z_o)\left(\frac{\partial \phi}{\partial z}\right)_o = 0,$$

and these may written in the standard form for the line,

$$\frac{x - x_o}{\begin{vmatrix} \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{vmatrix}_o} = \frac{y - y_o}{\begin{vmatrix} \frac{\partial f}{\partial z} & \frac{\partial f}{\partial x} \\ \frac{\partial \phi}{\partial z} & \frac{\partial \phi}{\partial x} \end{vmatrix}_o} = \frac{z - z_o}{\begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} \end{vmatrix}_o}. \tag{12}$$

The parametric equations of a curve

$$x = f_1(t), \quad y = f_2(t), \quad z = f_3(t), \tag{13}$$

yield for the equations of the tangent line

$$\begin{aligned} x - x_o &= (t - t_o)\left(\frac{dx}{dt}\right)_o, \\ y - y_o &= (t - t_o)\left(\frac{dy}{dt}\right)_o, \\ z - z_o &= (t - t_o)\left(\frac{dz}{dt}\right)_o, \end{aligned} \tag{14}$$

from which we have

$$\frac{x - x_o}{\left(\frac{dx}{dt}\right)_o} = \frac{y - y_o}{\left(\frac{dy}{dt}\right)_o} = \frac{z - z_o}{\left(\frac{dz}{dt}\right)_o}. \tag{15}$$

The equation of a **plane normal to a curve** may have the form

$$A(x - x_o) + B(y - y_o) + C(z - z_o) = 0, \tag{16}$$

where A, B and C are the denominators of (12) or of (15) as convenience may require.

2. **Extension of Taylor's Theorem.** Let us consider z defined as a function of t by means of the equations

$$z = f(x, y), \quad x = x_o + ht, \quad y = y_o + kt, \tag{17}$$

then we shall have

$$\frac{dz}{dt} = \frac{\partial}{\partial x} f(x, y) \frac{dx}{dt} + \frac{\partial}{\partial y} f(x, y) \frac{dy}{dt} = h \frac{\partial}{\partial x} f(x, y) + k \frac{\partial f}{\partial y} f(x, y).$$

It is convenient to indicate this symbolically

$$\frac{dz}{dt} = \left[h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right] f(x, y),$$

the bracketed expression denoting an operation to be applied to the function f(x, y). Again,

$$\frac{d^2z}{dt^2} = h^2 \frac{\partial^2}{\partial x^2} f(x, y) + hk \frac{\partial^2}{\partial y \partial x} f(x, y) + hk \frac{\partial^2}{\partial x \partial y} f(x, y) + k^2 \frac{\partial^2}{\partial y^2} f(x, y),$$

and assuming that all functions involved are continuous so that the order of taking derivatives is immaterial, we have

$$\frac{\partial^2}{\partial y \, \partial x} f(x, y) = \frac{\partial^2}{\partial x \, \partial y} f(x, y),$$

and the result may be written symbolically

$$\frac{d^2z}{dt^2} = \left[h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right]^2 f(x, y).$$

In the same way we should have for a derivative of any order,

$$\frac{d^n z}{dt^n} = \left[h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right]^n f(x, y), \tag{18}$$

indicating that the operations to be applied to f(x, y) are properly denoted by the terms of the binomial expansion

$$\left[h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right]^n = h^n \frac{\partial^n}{\partial x^n} + nh^{n-1} k \frac{\partial^n}{\partial y \, \partial x^{n-1}}$$

$$+ \frac{n(n-1)}{2!} h^{n-2} k^2 \frac{\partial^n}{\partial y^2 \, \partial x^{n-2}} + \cdots$$

Now we may apply Taylor's theorem to the function z of the independent variable t,

$$z = z_0 + t \left(\frac{dz}{dt} \right)_0 + \frac{t^2}{2!} \left(\frac{d^2z}{dt^2} \right)_0 + \frac{t^3}{3!} \left(\frac{d^3z}{dt^3} \right)_0 + \cdot + \frac{t^n}{n!} \left(\frac{d^n z}{dt^n} \right)_\theta \tag{19}$$

Here on the right, z and all successive derivatives of z are evaluated at t = o, excepting the nth derivative, involved in the remainder term, which is evaluated at θ, some value between 0 and t.

Substituting in (19) from previous equations we have

$$f(x, y) = f(x_o, y_o) + t\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right]f(x_o y_o) + \cdots + \frac{t^n}{n!}\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right]^n f(x_o + \theta h, y_o + \theta k), \tag{20}$$

and this is valid for all values of x_o, y_o, h, k and t such that the functions involved are continuous from x_o to $x_o + th$ and from y_o to $y_o + tk$. Since $th = x - x_o$ and $tk = y - y_o$

$$f(x_o y) = f(x_o, y_o) + \left[(x - x_o)\frac{\partial}{\partial x} + (y - y_o)\frac{\partial}{\partial y}\right]f(x_o, y_o) + \cdots + \frac{1}{n!}\left[(x - x_o)\frac{\partial}{\partial x} + (y - y_o)\frac{\partial}{\partial y}\right]^n f(c_1, c_2) \tag{21}$$

where c_1 is between x_o and x and c_2 between y_o and y. Another form of the same theorem is obtained if in (20) we put t = 1.

$$f(x_o + h, y_o + k) = f(x_o, y_o) + \left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right]f(x_o, y_o) + \cdots + \frac{1}{n!}\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right]^n f(x_o + \theta h, y_o + \theta k). \tag{22}$$

In (22) the subscripts of x_o and y_o may be omitted since the notations x and y no longer appear.

If, as n increases indefinitely in (21) or (22) the remainder term approaches zero, we obtain an infinite series in powers of $x - x_o$ and $y - y_o$, or of h and k, which represents perfectly the function on the left. Whether or not the remainder approaches zero will usually depend on x_o, y_o and h and k. For suitably chosen values good approximations may be obtained by omitting all derivatives of order higher than the first. This is, in fact, substituting the tangent plane at the point (x_o, y_o, z_o) for the surface, and the result is good in the neighborhood of this point, i.e., if h and k are small.

As an illustration, the range of an unresisted projectile is given by the equation

$$R = \frac{V^2 \sin 2\alpha}{g},$$

where V is the initial velocity, α the angle of elevation and g the acceleration due to gravity. We have

$$R - R_o = \frac{2V_o \sin 2\alpha_o}{g}(V - V_o) + \frac{2V_o^2 \cos 2\alpha_o}{g}(\alpha - \alpha_o)$$

neglecting the powers of $V - V_o$ and $\alpha - \alpha_o$ after the first. This is an effective means of computing the change $R - R_0$ that is produced in the range by small changes $V - V_o$ and $\alpha - \alpha_o$ in the initial velocity and elevation.

The method we have used here to extend Taylor's theorem to two independent variables may be used for any number of variables. We state here, for three independent variables, the form analogous to (22).

$$f(x + h, y + k, z + m) = f(x, y, z) + \left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} + m\frac{\partial}{\partial z}\right] f(x, y, z)$$

$$+ \frac{1}{2!}\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} + m\frac{\partial}{\partial z}\right]^2 f(x, y, z) + \cdots \quad (23)$$

$$+ \frac{1}{n!}\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} + m\frac{\partial}{\partial z}\right]^n f(x + \theta h, y + \theta k, z + \theta m),$$

where $0 < \theta < 1$.

3. **Maxima and Minima.** Let us now inquire under what conditions the quantity $f(x + h, y + k, z + m) - f(x, y, z) = Q$ shall be negative or zero for all small values of h, k, and m. Under these conditions the function is said to be at a **maximum** at (x, y, z). If we take $n = 1$ in (23) above we obtain

$$Q = h\frac{\partial}{\partial x} f(x + \theta h, y + \theta k, z + \theta m)$$

$$+ k\frac{\partial}{\partial y} f(x + \theta h, y + \theta k, z + \theta m)$$

$$+ m\frac{\partial}{\partial z} f(x + \theta h, y + \theta k, z + \theta m).$$

Now take $k = 0$ and $m = 0$, then if Q is to be negative the two factors, h and the partial derivative with respect to x, must be of opposite signs, and as h passes through zero from negative to positive the other factor must pass through zero from positive to negative. Thinking of this factor as a function of h, and applying the ideas of

Art. 37, we find that the derivative with respect to h of this function is negative or zero, that the first successive derivative which is not zero must be of odd order, and negative. Now derivatives with respect to h have the same signs as derivatives with respect to x, and the argument here for h may be repeated for k and m, hence for a maximum at (x, y, z) we must have

$$\frac{\partial}{\partial x} f(x, y, z) = 0, \frac{\partial}{\partial y} f(x, y, z) = 0, \frac{\partial}{\partial z} f(x, y, z) = 0, \quad (24)$$

$$\frac{\partial^p}{\partial x^p} f(x, y, z) < 0, \frac{\partial^q}{\partial y^q} f(x, y, z) < 0, \frac{\partial^r}{\partial z^r} f(x, y, z) < 0, \quad (25)$$

where p is the order of the first one of the successive partial derivatives with respect to x that is not zero at (x, y, z), and p must be even. The integers q and r are similarly defined and must be even.

If the quantity Q is positive for all small values of h, k and m, when not all are zero, the function f(x, y, z) is said to be at a **minimum** at (x, y, z). The conditions (24) and (25) with the inequality reversed, must be satisfied for a minimum, as may be shown by a similar argument.

The discussion here given for a function of three variables may be applied to any number of variables. In the case of two variables, the condition that the first partial derivatives shall be zero is equivalent to the condition that the tangent plane shall be horizontal. For any number of variables, if we suppose all variables but one, say x, to hold fixed values, then the condition that the function shall be a maximum for x is, according to Art. 38, exactly the same as the conditions here involving partial derivatives with respect to x. The conditions (24) and (25), may be regarded as demanding that f(x, y, z) shall be a maximum for x, regarding y and z constant, a maximum for y regarding z and x constant and a maximum for z regarding x and y constant.

While the conditions here stated for a maximum or a minimum are necessary, they are not sufficient, that is, they are not conclusive. In practice, whether the function under consideration is expressed explicitly in terms of the independent variables or is determined by two or more relations involving auxiliary variables, the determination of a maximum or minimum is usually accomplished by application of conditions (24), the practical conditions of the problem making further tests unnecessary. In the case of implicit functions or auxiliary variables the equations of (24) are taken together with the given equations to determine the position of the maximum or minimum. Of

course conditions (25) may always be used to discriminate between a maximum and a minimum, but these conditions may fail to detect the absence of an extreme of either sort.

We may state here, that for any function, z, of two variables x and y, an extreme certainly exists at a point for which (24) and (25) are satisfied and also

$$\frac{\partial^2 z}{\partial x^2}\frac{\partial^2 z}{\partial y^2} - \left[\frac{\partial^2 z}{\partial x\,\partial y}\right]^2 > 0. \tag{26}$$

If the inequality is reversed, no extreme exists; if the expression is zero, the test is not conclusive.

In case it is desired, for any given function, to carry the investigation beyond the conditions stated here, two courses are suggested. First, a geometrical study of the function by the methods of the two previous articles may yield results. Second, an expression of Q by means of (23), where n is taken greater than unity, or, in the case of integral algebraic functions, where the expansion is carried out completely, may enable us to determine whether Q remains always positive or always negative for small values of h, k and m.

As an example, let us test for extremes the function

$$z = x^3 + y^3 - 6xy.$$

Conditions (24) give

$$\frac{\partial z}{\partial x} = 3x^2 - 6y = 0, \frac{\partial z}{\partial y} = 3y^2 - 6x = 0.$$

The real solutions of these equations are

$$x = 0, y = 0, \text{ and } x = 2, y = 2.$$

Now we have

$$\frac{\partial^2 z}{\partial x^2} = 6x, \frac{\partial^3 z}{\partial x^3} = 6, \frac{\partial^2 z}{\partial y^2} = 6y, \frac{\partial^3 z}{\partial y^3} = 6.$$

For the point (0, 0) the second derivatives are zero, and the third derivatives are not zero, therefore, the requirement for even order in (25) is not satisfied, and (0, 0) is not an extreme. At (2, 2) the second derivatives are positive, hence the point may give a minimum. Here $\frac{\partial^2 z}{\partial x\,\partial y} = -6$, so that, by (26) the minimum is established at (22),

the value being $z = -8$. If we apply equation (22) here, taking $n = 2$, we have for Q the expression

$$\frac{1}{2}\left[h^2\frac{\partial^2 z}{\partial x^2} + 2hk\frac{\partial^2 z}{\partial x \partial y} + k^2\frac{\partial^2 z}{\partial y^2}\right]$$

evaluated at $x = 2 + \theta h$, $y = 2 + \theta k$, giving the quantity

$$Q = \frac{1}{2}h^2(12 + 6\theta h) - 12hk + k^2(12 + 6\theta k)$$

and this reduces to

$$Q = (h^2 - hk + k^2)(6 + 3\theta h + 3\theta k).$$

Of these two factors the former is positive for all values of h and k and the latter is positive for small values of h and k, so that this test is sufficient to establish the minimum at (2, 2).

If the expansion (22) is carried out completely we have

$$z = -8 + 6(h^2 - hk + k^2) + h^3 + k^3$$

which gives

$$Q = z + 8 = (h^2 - hk + k^2)(6 + h + k),$$

showing that $\theta = \frac{1}{3}$ in the expression above, and establishing, for the third time, the minimum at the point (2, 2).

As a second example consider

$$z = x^4 - 3x^2y + 2y^2.$$

We have

$$\frac{\partial z}{\partial x} = 4x^3 - 6xy, \qquad \frac{\partial z}{\partial y} = -3x^2 + 4y,$$

$$\frac{\partial^2 z}{\partial x^2} = 12x^2 - 6y, \qquad \frac{\partial^2 z}{\partial y^2} = 4.$$

$$\frac{\partial^3 z}{\partial x^3} = 24x,$$

$$\frac{\partial^4 z}{\partial x^4} = 24.$$

The only critical point is at (0, 0), and at this point conditions (25), with inequalities reversed, are satisfied, suggesting a minimum. Now $\frac{\partial^2 z}{\partial x \partial y} = -6x$, so that (26) gives, at (0, 0),

$$\frac{\partial^2 z}{\partial x^2}\frac{\partial^2 z}{\partial y^2} - \left[\frac{\partial^2 z}{\partial x\,\partial y}\right]^2 = 0 \cdot 4 - 0 = 0,$$

so that (26) is not conclusive.

In this case the function z may be factored, $z = (x^2 - y)(x^2 - 2y)$, and if the two factors are equated to zero, the loci in the xy-plane are parabolas having a common vertex at the origin. Between these parabolas z is negative, elsewhere positive, so that z can have no extreme at the origin.

The equation (22), for $n = 2$, gives in this case, evaluating the remainder term at $x = \theta h$, $y = \theta k$,

$$Q = \frac{1}{2}[h^2(12\theta^2 h^2 - 6\theta k) - 12hk + 4k^2]$$

$$= h^2(6\theta^2 h^2 - 3\theta k) - k(6h - 2k).$$

Here either term may be positive or negative according to the relations between h and k, however small these quantities may be, showing again that no extreme exists.

If the expansion (22) were carried out completely in this case we should merely be replacing x and y by h and k in the equation for z, and the factored form shown above would be effective in showing that no extreme exists.

EXERCISES LII

1. Write the equation of a plane tangent to the sphere $x^2 + y^2 + z^2 = 169$ at the point (3, −4, 12). Show that the radius from the origin to this point is normal to this plane.

2. Write the equation of tangent plane and normal line to the surface. $x = u^3 + v^2$, $y = u^2 - v$, $z = u + v^2$ at the point where $u = 1$ and $v = 2$.

3. Equations of two curves on the surface in example (2) are found by putting first $u = 1$, and then $v = 2$. Find tangent lines to these curves at $v = 2$ and $u = 1$, and show that they lie in the tangent plane and are perpendicular to the normal line found in (2).

4. Write the equations of normal plane and tangent line to the curve

given by the equations $x + z - 2xy = 12$, $y + z - 2xz = -8$, the point of contact being $(2, -2, 2)$.

5. By the extension of Taylor's theorem write an approximation for $\sin(x + y)$ in powers of x and y, up to the third powers. Check the result by putting $y = x$ and comparing with the series for $\sin 2x$.

6. The area of a triangle being given by the formula $A = \frac{1}{2}ab \sin C$ and the measurements being $a = 1000$ ft., $b = 600$ ft., $C = 42°$, find the limit of error in A if a and b are correct to one-half of one per cent and C is correct within a tenth of a degree.

7. Find the equations of the tangent line and normal plane to the helix $x = 5\cos\theta$, $y = 5\sin\theta$, $z = \frac{1}{2}\theta$, at $\theta = \theta_0$. Show that the normal plane contains the point $\left(0, 0, \frac{1}{2}\theta_0\right)$.

8. Show that the tangent plane to the sphere $x^2 + y^2 + z^2 + 2ax + 2by + 2cz + d = 0$ at the point (x_1, y_1, z_1) is $x_1x + y_1y + z_1z + a(x + x_1) + b(y + y_1) + c(z + z_1) + d = 0$.

9. A number is to be divided into three parts so that their product is a maximum. Show that the three parts must be equal.

10. Show that the surface of a rectangular parallelopiped of given volume is least when the solid is a cube.

11. Determine the maxima and minima of z on the surface

$$x^2 + 2y^2 + 3z^2 - 2xy - 2yz = 2.$$

12. Show that the function

$$w = 2x^2 - 3y^2 + 5z^2$$

has neither maximum nor minimum.

13. The four vertices of a tetrahedron are (x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2), $x_3, y_3, z_3)$. Show that if (x, y, z) is so chosen that the sum of the squares of its distances from these vertices is a minimum, then $x = \frac{1}{4}(x_0 + x_1 + x_2 + x_3)$, $y = \frac{1}{4}(y_0 + y_1 + y_2 + y_3)$ and $z = \frac{1}{4}(z_0 + z_1 + z_2 + z_3)$.

14. Express x^2y in powers of $x - 2$ and $y + 1$ by use of equation (21).

15. Write the expansion for e^{xy} in powers of x and y up to and including the third degree terms. Express the remainder term after the second degree terms.

16. Show that the polynomial of the second degree which is the

proper approximation for the function $\log (\sin x + \cos y)$ when x and y are small is $x - \frac{1}{2}(x^2 + y^2)$.

17. Write the second degree approximation for

$$\frac{y}{2}\sqrt{y^2 - x^2} + \frac{y^2}{2}\sin^{-1}\frac{x}{y},$$

when x is small and y is near unity.

Art. 56. Certain Properties of Plane Curves

1. **Singular points.** In previous articles we have seen how the study of surfaces is facilitated by consideration of the curves in which they cut certain planes. Now the discussion of certain properties of plane curves gains clearness by associating them with properties of surfaces on which the curves lie. The curve

$$f(x, y) = 0 \tag{1}$$

is a plane section of the surface

$$z = f(x, y). \tag{2}$$

We have

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}, \tag{3}$$

and if z is regarded as a constant, positive, negative or zero, then (2) is the equation of a plane section of the surface and (3) gives

$$\frac{dy}{dx} = -\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial y}}. \tag{4}$$

Suppose $z = z_o$, then if $f(x_o y_o) = z_o$, the expression (4), evaluated at (x_o, y_o) will in general give the slope, $\frac{dy}{dx}$, at that point on the curve $f(x, y) = z_o$. It may happen, however, that for the point (x_o, y_o) we have $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$. In this case (4) is indeterminate, and does not yield a definite value for the slope. Such a point (x_o, y_o) is called a **singular point** of the curve $f(x, y) = z_o$. It will be noticed that such

a point is a critical point in the investigation of maxima and minima for z by the previous article.

To find the singular points of (1) we equate the partial derivatives to zero and solve for x and y. Let the values found be x_o and y_o, then if $f(x_o, y_o) = 0$, the point is a singular point of (1), but if $f(x_o, y_o) = z \neq 0$, then the point is not on (1), but pertains to the curve $f(x, y) = z_o$.

Singular points are of several kinds. In the example $z = x^3 + y^3 - 6xy$ considered in the last article we found critical points at (0, 0) and (2, 2). The point (0, 0) is a singular point on the curve $x^3 + y^3 - 6xy = 0$, of the kind called a **double point,** since the curve crosses itself at the origin, as may be seen by plotting. The point (2, 2) is a singular point of the curve $x^3 + y^3 - 6xy = -8$, and is an **isolated point** or **conjugate point,** since no other point of the curve is near this point. It may be seen that a maximum or minimum point for z on the surface $z = f(x, y)$ is an isolated point for the corresponding curve $f(x, y) = z_o$. In the example $z = (x^2 - y)(x^2 - 2y)$ the critical point was (0, 0), and consists of a **contact point** between two branches of the curve $(x^2 - y)(x^2 - 2y) = 0$, sometimes called a **tacnode.**

For the curve $x^2 - y^3 = 0$, we determine the singular point by the equations

$$\frac{\partial f}{\partial x} = 2x = 0, \qquad \frac{\partial f}{\partial y} = -3y^2 = 0,$$

from which $x = 0$, $y = 0$, and, since the origin is on the curve it is a singular point. Plotting of the curve reveals that fact that the curve is tangent to the y-axis on both sides at the origin but is not below the x-axis. Where two branches of a curve thus meet at a common tangent, but terminate at this common point the curve is said to have a **cusp.** If the two branches meet and terminate at the common point but have not a common tangent, then the curve has a corner, called a **salient point.**

A complete discussion of the means for discriminating between the various kinds of singular points may not be undertaken here, but once the point is located, plotting and considering the associated surface should be sufficient for most practical purposes.

2. **Envelopes.** Consider a family of curves given by the equation

$$f(x, y, k) = 0, \tag{5}$$

where k is the parameter of the family. If we substitute for x and y the coordinates of a particular point (x_o, y_o) then (5) becomes an

equation in which k is the unknown, and determines those values of k for which the curve will pass through (x_o, y_o). The number of distinct curves of the family that pass through the point depends on the number of real distinct valves of k, and the xy-plane may be divided into regions according to the number of curves passing through each point of the region. In Figure 107 the curve AB divides the plane into two regions; on one side, as at P_1 and P_2, two distinct curves pass through each point, while on the other side, as at P_4, no curve passes through the point. On the boundary between two such regions, as at P_3, two values of k must coincide.

The condition that two values of k given by (5) shall coincide is that this equation shall have a root in common with the equation found

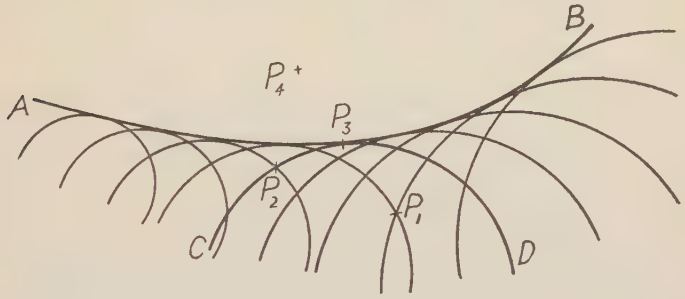


Fig. 107

by differentiating with respect to k, as shown in Art. 37. The two equations

$$f(x, y, k) = 0,$$

$$\frac{\partial}{\partial k} f(x, y, k) = 0, \tag{6}$$

therefore must be satisfied by the same value of k if x and y are so chosen that the point (x, y) lies on such a boundary curve, as AB. Elimination of k between equations (6) will give an equation in x and y the locus of which must include the boundary curve, called the **envelope** of the family of curves (5).

We may regard the two equations (6) as determining x and y as functions of a parameter k, so that the locus is given in parametric form. Suppose that x_1 and y_1 are values corresponding to the value k_1 of the parameter. Let P_3 represent the point, then CD is the curve

given by (5) when $k = k_1$. We have, for these particular values of the variables,

$$f(x_1, y_1, k_1) = 0, \quad \frac{\partial}{\partial k} f(x_1, y_1, k_1) = 0. \tag{7}$$

To find the slope of CD at P_3 we use equation (5), with $k = k_1$, and find

$$\left[\frac{\partial f}{\partial x}\right]_1 + \left[\frac{\partial f}{\partial y}\right]_1 \frac{dy}{dx} = 0, \tag{8}$$

where the subscript indicates that the expression has been evaluated at $x = x_1$, $y = y_1$ and $k = k_1$. To find the slope of AB at P_3 we use equations (6), regarding x and y as functions of k. We must have

$$\left[\frac{\partial f}{\partial x}\right]_1 \cdot \frac{dx}{dk} + \left[\frac{\partial f}{\partial y}\right]_1 \cdot \frac{dy}{dk} + \left[\frac{\partial f}{\partial k}\right]_1 = 0. \tag{9}$$

Now the third term here is equal to zero by (7), therefore, if equation (8) yields a definite value for the slope of CD at P_3, equation (9) yields the same for the slope of AB at the same point. At every point of the envelope, therefore, there is a curve of the family having a common tangent with the envelope.

Sometimes the locus given by equations (6) contains a curve every point of which is a singular point of one of the curves given by (5). For points on such a curve equation (8) does not yield a definite value for the slope, and such a curve is not a part of the envelope.

As an illustration consider

$$f(x, y, k) = x^2 + (y - k)^2 - 16 = 0,$$

$$\frac{\partial}{\partial k} f(x, y, k) = -2(y - k) = 0.$$

Eliminating k we find $x = \pm 4$ for the envelope. Here the curves are circles of radius 4 with centers on the y-axis, and the envelope consists of the two vertical lines tangent to all of the circles.

As another illustration consider

$$f(x, y, k) = x^3 - (y - k)^2 = 0,$$

$$\frac{\partial}{\partial k} (x, y, k) = 2(y - k) = 0.$$

Here elimination of k gives $x = 0$, the equation of the y axis. This

is not an envelope, however, for it touches the curve $x^3 - (y - k)^2 = 0$ at $(0, k)$ and we have

$$\frac{\partial f}{\partial x} = 3x^2, \quad \frac{\partial f}{\partial y} = -2(y - k),$$

both of which are zero at $(0, k)$ showing that the locus is a locus of singular points. The curves are, in fact, semi-cubical parabolas, with cusps on the y-axis, the tangent at the cusp being the horizontal line through the point.

3. **Evolutes.** Given a curve

$$f(x, y) = 0. \tag{10}$$

we may write the equation of the tangent at any point (x_o, y_o) of the curve in the form

$$(x - x_o)\left[\frac{\partial f}{\partial x}\right]_o + (y - y_o)\left[\frac{\partial f}{\partial y}\right]_o = 0, \tag{11}$$

and the equation of the normal at the same point is

$$(x - x_o)\left[\frac{\partial f}{\partial y}\right]_o - (y - y_o)\left[\frac{\partial f}{\partial x}\right]_o = 0. \tag{12}$$

Now if we regard y_o as a function of x_o given by the equation $f(x_o, y_o) = 0$, then (12) is a family of lines with one parameter, x_o. To find the envelope of this family we take the derivative of (12) with respect to x_o, treating x and y as constants but remembering that y_o is a function of x_o, then eliminate x_o and y_o, obtaining the equation of the envelope in terms of x and y. The envelope of the family of normals to a curve is called the **evolute** of the curve.

To illustrate, consider the parabola

$$y - x^2 = 0. \tag{13}$$

The normal at (x_o, y_o) is

$$(x - x_o) + (y - y_o)\,2x_o = 0,$$

and since $y_o = x^2_o$, this reduces to

$$x + (2y - 1)\,x_o - 2x_o^3 = 0. \tag{14}$$

The partial derivative with respect to the parameter x_o is now equated to zero, giving

$$2y - 1 - 6x_o^2 = 0. \tag{15}$$

Eliminating x_0 between (14) and (15) we find the evolute in the form

$$27x^2 = 16\left(y - \frac{1}{2}\right)^3.$$

In practice the elimination of x_0 and y_0 is likely to be laborious, and a form of the equation of the normal involving only one parameter can be found more advantageously than by a direct use of (12). Instances are shown in Art. 35, where families of normals to the conic sections are shown involving the slope as a parameter.

If the curve (10) is given in parametric form

$$x = \theta(t), \quad y = \phi(t), \tag{16}$$

the normal at $t = t_0$ is written

$$[x - \theta(t_0)]\,\theta'(t_0) + [y - \phi(t_0)]\,\phi'(t_0) = 0, \tag{17}$$

a form involving only one parameter, t_0. As an illustration of this take the ellipse

$$x = a\cos t, \quad y = b\sin t \tag{18}$$

then the normal at any value t_0 is

$$-(x - a\cos t_0)\,a\sin t_0 + (y - b\sin t_0)\,b\cos t_0 = 0,$$

or reducing,

$$ax\sin t_0 - by\cos t_0 = (a^2 - b^2)\sin t_0\cos t_0. \tag{19}$$

Differentiating with respect to t_0 we obtain

$$ax\cos t_0 + by\sin t_0 = (a^2 - b^2)(\cos^2 t_0 - \sin^2 t_0). \tag{20}$$

Multiplying (19) by $\sin t_0$ and (20) by $\cos t_0$ and adding we have

$$ax = (a^2 - b^2)\cos^3 t_0,$$

and multiplying (19) by $\cos t_0$ and (20) by $\sin t_0$ and subtracting the former we have

$$by = -(a^2 - b^2)\sin^3 t_0.$$

These last two equations give the evolute in convenient parametric form. The standard rectangular equation is found by taking the $\frac{2}{3}$ power of both sides of both equations and adding:

$$a^{\frac{2}{3}}x^{\frac{2}{3}} + b^{\frac{2}{3}}y^{\frac{2}{3}} = (a^2 - b^2)^{\frac{2}{3}}.$$

4. **Asymptotes.** In previous chapters we have seen that a curve has a vertical asymptote $x = a$ if y becomes infinite as x approaches a and that it has a horizontal asymptote $y = b$ if x becomes infinite as y approaches b. Also we have seen that if the product of linear factors is equated to a constant then each factor equated to zero gives the equation of an asymptote. In the present section we discuss more general methods for finding oblique asymptotes.

Let equation (10) represent a curve, then (11) is the equation of a tangent at (x_o, y_o). The intercepts of (11) are

$$a = x_o + y_o \frac{\left[\frac{\partial f}{\partial y}\right]_o}{\left[\frac{\partial f}{\partial x}\right]_o}, \quad b = x_o \frac{\left[\frac{\partial f}{\partial x}\right]_o}{\left[\frac{\partial f}{\partial y}\right]_o} + y_o.$$

If a and b remain finite when x_o, and y_o become infinite, still subject to the condition $f(x_o, y_o) = 0$, then the line (11) becomes an asymptote.

To illustrate, consider the curve

$$y^3 = 6x^2 + x^3.$$

The intercepts of the tangent are, by (11)

$$a = x_o - \frac{y_o^3}{4x_o + x_o^2}, \quad b = -\frac{4x_o^2 + x_o^3}{y_o^2} + y_o.$$

Since $y_o^3 = 6x_o^2 + x_o^3$, these reduce to the expressions

$$a = -\frac{2x_o^2}{4x_o + x_o^2}, \quad b = \frac{2x_o^2}{[6x_o^2 + x_o^3]^{\frac{2}{3}}}.$$

The limits of these as x_o increases indefinitely are -2 and $+2$, so that the asymptote is

$$x - y + 2 = 0.$$

Another method for finding oblique asymptotes is often more convenient than that above. A line $y = mx + b$ may intersect the curve (10) in two or more points, and to find the abscissas of such points we should eliminate y, obtaining an equation for x,

$$f(x, mx + b) = 0.$$

If m and b are so chosen that the coefficients of the two highest powers of x are zero, then the line meets the curve at two infinite points, and

is an asymptote. Applying this method to the example used above we have

$$m^3x^3 + 3bm^2x^2 + 3b^2mx + b^3 = 6x^2 + x^3,$$

and the coefficients of x^3 and x^2 are zero if $m^3 = 1$ and $3bm^2 = 6$, giving $m = 1$, $b = 2$, and the asymptote is $y = x + 2$.

EXERCISES LIII

1. Show that the curve $y^2 = x^3 - 4x^2$ has an isolated point at the origin.
2. Find the singular points for $y^5 + 2xy^2 = x^2 + y^4$.
3. Find the envelope of the family of lines

$$y = mx + \sqrt{1 + m^2}.$$

4. Find the envelope of the family of circles

$$x^2 + 4ax + y^2 + 3a^2 = 0.$$

5. Find the envelope of the family of lines

$$x \cos \alpha + y \sin \alpha = 5.$$

6. Find the evolute of the cissoid $y^2 = \dfrac{x^3}{2a - x}$.
7. Show that the evolute of a circle consists of a single point.
8. Using the method of Art. 39 show that the evolutes found in exercises 6 and 7 are the loci of the centers of curvature of the two curves.
9. Find asymptotes to the curve

$$x^2 - xy - 6y^2 + 2x + 4y + 12 = 0.$$

10. Find asymptotes to the curve $y^2 = \dfrac{x^3}{x - 2}$.
11. Show that the evolute of the cycloid

$$x = a(t - \sin t), \quad y = a(1 - \cos t)$$

is another cycloid given by the equations

$$x = a(t + \sin t), \quad y = -a(1 - \cos t),$$

which reduce to the given equations if t is increased by π, x by $-a\pi$ and y by 2a. Draw the curves.

12. Find the evolute of the curve

$$y = \sec\theta, \quad x = \tan\theta,$$

both as the locus of the center of curvature and as the envelope of the normal lines.

13. Discuss the following curves as to the singular points and asymptotes. Trace the curves, noting points of inflexion, and extremes:

(a) $y^2 = \dfrac{1}{x - a}$; (b) $y^2 = x^2 - x^4$; (c) $y = \dfrac{x}{(x - a)^2}$;

(d) $x^2y^2 - x^4 + y^2 + x^2 = 0$; (e) $xy^2 - x^3 - 2x^2 + 1 = 0$.

14. Tangents to the hyperbola $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ are given by the equation

$$y = mx \pm \sqrt{a^2m^2 - b^2},$$

where m is arbitrary. Show that the envelope of this family of tangents is the given hyperbola.

15. Find the evolute of the curve

$$x = a(\cos\theta + \theta\sin\theta),$$

$$y = a(\sin\theta - \theta\cos\theta).$$

Art. 57. Integrals in Several Variables

1. **Multiple integrals.** Consider a function of several variables, and for convenience let us fix attention on the case of three variables. Let the function be defined by the equation

$$w = f(x, y, z), \tag{1}$$

and regard x, y and z as the coordinates of a point in three dimensions. Let R denote a certain region in this three dimensional space, and consider R as made up of any number of differential elements of dimensions dx, dy and dz, together with certain fractional portions around the boundaries of R. The volume of the differential element is dx dy dz, and if the dimensions dx, dy and dz all approach zero, the sum of the volumes of all such elements within R will approach the volume of R as a limit. Now if the volume of each differential element is multiplied by the value of the function w at some point (x, y, z) within the element then the sum of such products

$$\Sigma\, f(x, y, z)\, dx\, dy\, dz \tag{2}$$

may approach a limit when the sum of the elementary volumes approaches the volume of R. If so, then the limit of the sum (2) is called the integral of f(x, y, z) over the region R. We may employ the notations

$$\int_R w \, d\, V = \lim \Sigma_R \, f(x, y, z) \, dx \, dy \, dz$$

$$= \iiint_R f(x, y, z) \, dx \, dy \, dz. \qquad (3)$$

There being three variables, such a quantity is called a **triple integral.** The idea is applicable to any number of variables, with suitable modification of the geometrical interpretation, the general term being **multiple integral.**

A **double integral** may be interpreted as pertaining to a portion of a plane, and taking $z = f(x, y)$ we may write

$$\int_R z \, d\, A = \lim \Sigma_R \, f(x, y) \, dx \, dy = \iint_R f(x, y) \, dx \, dy, \qquad (4)$$

where $d\, A = dx \, dy$ is the area of a differential element of the portion R of the plane. A double integral may be evaluated as an iterated integral in two variables, such as was treated in Art. 44. Thus we may write

$$\int_R z \, d\, A = \int_{x_1}^{x_2} \int_{\phi_1(x)}^{\phi_2(x)} f(x, y) \, dy \, dx, \qquad (5)$$

if the region R is bounded by the curves $y = \phi_1(x)$ and $y = \phi_2(x)$ and the lines $x = x_1$ and $x = x_2$. If R is not bounded by such loci, then it may be divided into several parts, each of which is so bounded, and the double integral may be given as the sum of several iterated integrals, each covering one of the divisions of R. This method of evaluating the double integral is based on taking the limit of the sum in (4) by first taking the limit of the sum of one column of elements, the length of the column being $\phi_2(x) - \phi_1(x)$ and the width dx, then taking the limit of the sum of all such columns from x_1 to x_2. Critical discussion of the equivalence of the double integral and the iterated integral is omitted here.

In similar manner the triple integral in (3) may be evaluated as an iterated integral in three variables. We have

$$\int_R w \, dv = \int_{x_1}^{x_2} \int_{\phi_1(x)}^{\phi_2(x)} \int_{\theta_1(x, y)}^{\theta_2(x, y)} f(x, y, z) \, dz \, dy \, dx, \qquad (6)$$

where the region R occupies the space between the surfaces $z = \theta_1$ (x, y) and $z = \theta_2$ (x, y) and between the cylindrical surfaces $y = \phi_1(x)$ and $y = \phi_2(x)$, and between the planes $x = x_1$ and $x = x_2$. If R may not be so bounded, then it may be subdivided into parts that are so bounded and the triple integral expressed as a sum of iterated integrals. Of course the variables may be interchanged in (6), as well as in (5), so that the integrations may be done in any order, according to convenience.

While the double and triple integrals are always capable of geometrical interpretation as indicated, the definitions are not essentially geometrical, and in many uses of the multiple integral the variables are not regarded as rectangular coordinates. The limit of the sum in (3) or (4) is a definite concept entirely apart from any geometrical interpretation, the geometrical language being merely a convenience in stating the quantitative relations that exist.

The applications of multiple integrals are so numerous and extensive that we must be content to indicate a few of the more common uses.

(a) **Volume in rectangular coordinates:** Let the volume be bounded as the region R in (6) then

$$V = \int_{x_1}^{x_2} \int_{\phi_1(x)}^{\phi_2(x)} [\theta_2(x, y) - \theta_1(x, y)]\, dy\, dx.$$

This same volume is given as a triple integral by (6) if f(x, y, z) is taken as unity throughout the region R. As an example consider the volume bounded by the planes $z = 0$ and $z = 1 + x + y$, and the parabolic cylinder $y = x^2$ and the plane $y = 4$. The triple integral becomes

$$\int_{-2}^{2} \int_{x^2}^{4} \int_{0}^{1+x+y} dz\, dy\, dx$$

Integrating with respect to z we have

$$\int_{-2}^{2} \int_{x^2}^{4} (1 + x + y)\, dy\, dx.$$

This is the double integral representation of the volume. Integrating now with respect to y, regarding x as constant, the new integrand is given by

$$\left[y + xy + \frac{y^2}{2}\right]_{x^2}^{4} = 4 + 4x + 8 - \left(x^2 + x^3 + \frac{x^4}{2}\right),$$

so that the integral becomes

$$\int_{-2}^{2} \left(12 + 4x - x^2 - x^3 - \frac{x^4}{2}\right) dx = 48 - \frac{16}{3} - \frac{32}{5} = 36.27.$$

b. **Volume in cylindrical coordinates.** Replace the rectangular coordinates (x, y, z) by the coordinates (ρ, θ, z) where $\rho \cos \theta = x$ and $\rho \sin \theta = y$. The three dimensions of the element of volume are $d\rho$, $\rho d\theta$ and dz, and volume is given by the triple integral

$$\int_R dV = \iiint_R \rho d\rho \, d\theta \, dz.$$

This may be evaluated as an iterated integral in any order. For example, find the volume of a sphere of radius 5. We write

$$V = \int_0^5 \int_0^{2\pi} \int_{\sqrt{25-\rho^2}}^{\sqrt{25-\rho^2}} \rho \, dz \, d\theta \, d\rho,$$

and integrating with respect to z we have

$$V = \int_0^5 \int_0^{2\pi} 2\rho \sqrt{25 - \rho^2} \, d\theta \, d\rho,$$

then integrating with respect to θ

$$V = \int_0^5 4\pi\rho \sqrt{25 - \rho^2} \, d\rho$$

whence

$$V = \left[-\frac{4}{3} \pi (25 - \rho^2)^{\frac{3}{2}} \right]_0^5 = \frac{500}{3} \pi.$$

(c) **Volume in spherical coordinates.** Take as coordinates of a point P the quantities ρ, θ and ϕ, where ρ is the distance from the origin, OP, θ is the angle between the xz-plane and the plane containing OP and the z-axis, and ϕ is the angle between OP and the z-axis. Thus θ may be regarded as longitude and ϕ as co-latitude on a sphere of radius ρ, in which the z-axis marks the poles and the xz-plane marks the prime meridian. We have the relations

$$x = \rho \sin \phi \cos \theta, \quad y = \rho \sin \phi \sin \theta, \quad z = \rho \cos \phi,$$

and we may take as three dimensions of a differential element $d\rho$, $\rho \sin \phi \, d\theta$ and $\rho d\phi$, so that volume is given by the triple integral

$$\int_R dV = \iiint_R \rho^2 \sin \phi \, d\rho \, d\theta \, d\phi.$$

If we apply this to a right circular cone, taking vertex at origin and axis of cone along the z-axis, we have

$$V = \int_0^{\alpha} \int_0^{2\pi} \int_0^{h \sec \phi} \rho^2 \sin \phi \, d\rho \, d\theta \, d\phi,$$

where α is the angle between axis and element of cone and h is altitude of cone. The steps in the integration are:

$$V = \int_0^{\alpha} \int_0^{2\pi} \frac{1}{3} h^3 \sec^3 \phi \sin \phi \, d\theta \, d\phi$$

$$= \int_0^{\alpha} \frac{2\pi}{3} h^3 \sec^3 \phi \sin \phi \, d\phi$$

$$= \frac{\pi}{3} h^3 (\sec^2 \alpha - 1) = \frac{\pi}{3} h^3 \tan^2 \alpha.$$

(d) **Moments and mean values.** If in the integrals for volume in (a), (b) and (c) above we insert an additional factor which represents, in terms of the coordinates employed, the distance of a point in the elementary volume from some fixed point, line or plane, then the resulting integral gives the **first moment** of the whole volume with respect to the given fixed point, line or plane. If this moment is divided by the volume, the quotient is the **mean value of the distance** for the volume considered. The mean values of distances from the three coordinate planes are the three rectangular coordinates of the center of gravity of the volume, which we may denote by x_m, y_m, z_m.

If the additional factor considered above is replaced by the square of distance the resulting integral gives the **second moment** or **moment of inertia** of the volume with respect to the fixed point, line or plane, and when this is divided by the volume the quotient is the mean value of the squared distance, and the square root of this mean value is the **radius of gyration** of the volume with respect to the given point, line or plane.

These notions of moment and mean value may be extended to any power of the distance, as in Art. 42. Insertion of the additional factor in the integrand does not involve any change in the limits of integration for a given volume, but the form of the new factor as a function of the variables may influence the choice of the system of coordinates and also the choice of order of integration.

(e) **Distributed quantities.** If a quantity is distributed throughout a volume, the "intensity" or "amount per unit volume" of the quantity may vary from point to point, and it may be expressible as a function of the coordinates of the point. If this intensity function is inserted as an additional factor in the integral for volume in (a), (b) or (c) then the resulting integral is the **total amount** of the quantity in the volume considered. If this total amount is divided by the volume the quotient is the **mean intensity** of the quantity in the given volume.

If this intensity factor and also a distance factor as in (d) are inserted in the integral for volume, then the result is a **moment of the total amount** of the quantity with respect to the point, line or plane from which distance is measured. If this moment is divided by the total amount the quotient is the mean distance of the total amount from the given point, line or plane. By using the square of the distance we obtain the second moment of the total amount, and the radius of gyration.

As an example let us consider the radius of gyration, with respect to a diameter, of a sphere whose density varies as the square of the distance from its center. Let the radius of the sphere be r, and take the center as origin and the given diameter as z-axis. In rectangular coordinates the intensity factor is $a(x^2 + y^2 + z^2)$, where a is constant, and the squared distance factor is $x^2 + y^2$, so that the total mass, M, of the sphere is given by the integral

$$M = 8 \int_0^r \int_0^{\sqrt{r^2 - x^2}} \int_0^{\sqrt{r^2 - x^2 - y^2}} a(x^2 + y^2 + z^2)\, dz\, dy\, dx,$$

where, on account of symmetry, we take 8 times the mass of the first octant. The integral for the moment of inertia I is like that for M but contains the additional factor $x^2 + y^2$ in the integrand. If $k^2M = I$, then k is the required radius of gyration.

In cylindrical coordinates the intensity factor is $a(\rho^2 + z^2)$ and the squared distance factor is ρ^2, and we have

$$I = 2 \int_0^r \int_0^{2\pi} \int_0^{\sqrt{r^2 - \rho^2}} a(\rho^2 + z^2)\, \rho^3\, dz\, d\theta\, d\rho,$$

taking twice the integral for the upper half of the sphere on account of symmetry. The integral for M is like that for I, but divided by the factor ρ^2.

In spherical coordinates the intensity factor is a ρ^2 and the squared distance factor is $\rho^2 \sin^2 \phi$, so that

$$M = \int_0^{\pi} \int_0^{r} \int_0^{2\pi} a\rho^4 \sin \phi \, d\theta \, d\rho \, d\phi,$$

$$I = \int_0^{\pi} \int_0^{r} \int_0^{2\pi} a\rho^6 \sin^3 \phi \, d\theta \, d\rho \, d\phi.$$

It is evident that, for the particular problem under consideration, these last integrals are more convenient than those in rectangular or cylindrical coordinates.

(f) **Curved surfaces.** Let a surface be given by an equation of the form $f(x, y, z) = 0$, and consider the area of a certain portion, S, of this surface. Let the projection of S on the xy-plane be R, and take the element of area of R as dx dy. The corresponding differential element of area on S is sec γ dx dy, where γ is the angle between the xy-plane and the tangent plane to S at some point in the elementary area, or, what is the same thing, the angle between the z-axis and the normal line to the surface. From Arts. 53 and 55 we see that

$$\sec \gamma = \frac{1}{\dfrac{\partial f}{\partial z}} \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2},$$

so that the area of S is given by

$$\int_R \sec \gamma \, dA = \iint_R \frac{1}{\dfrac{\partial f}{\partial z}} \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2} \, dx \, dy.$$

Moments, mean values and distributed quantities in relation to curved surfaces may be treated by inserting the appropriate distance and intensity factors in the integrand.

In evaluating this type of integral as an iterated integral it must be remembered that $f(x, y, z) = 0$ for all points on the surface, so that if z occurs in the integrand it must be expressed in terms of x and y before integrating.

2. Line Integrals. Consider the integral

$$\int_{x_1}^{x_2} P(x, y) \, dx, \tag{7}$$

where y is a function of x given by the equation

$$y = f(x). \tag{8}$$

This integral may be evaluated as a simple integral after substituting for y its value in terms of x. But let us fix attention on the significance of the integral in relation to the function P(x, y) and the curve $y = f(x)$. We see that the differential dx is multiplied by the value of P(x, y) at a point (x, y) on the curve $y = f(x)$, the abscissa of the point being an x within the interval covered by dx. The function P(x, y) assigns an "intensity" to each point (x, y) in the plane, and the curve $y = f(x)$ selects a series of points to be associated with the series of dx's. The integral is known as the **line integral** of P(x, y) along the curve $y = f(x)$ from where $x = x_1$ to where $x = x_2$. If y is a multiple valued function of x it is convenient to express the line integral as a sum of definite integrals, each covering a part of the curve along which y is single valued.

For example, let $P(x, y) = xy - y^2$ and let the curve have the equation $y^2 = 4x - 8$, and integrate from the point $(3, -2)$ to the point $(6, 4)$. The curve is a parabola with vertex at $(2, 0)$, with axis horizontal, so that $(3, -2)$ is on the lower branch and $(6, 4)$ on the upper branch. On the lower branch $y = -2\sqrt{x-2}$, and on the upper branch $y = 2\sqrt{x-2}$. The line integral is

$$\int_{3,-2}^{6,4} (xy - y^2)\,dx = \int_3^2 (-2x\sqrt{x-2} - 4x + 8)\,dx + \int_2^6 (2x\sqrt{x-2} - 4x + 8)\,dx.$$

Similarly we may have the line integral

$$\int_{y_1}^{y_2} Q(x, y)\,dy \tag{9}$$

where x is to be expressed as a function of y by the equation of the curve along which the integral is taken. If (7) and (9) are taken along the same curve the sum of the two is a line integral

$$\int_{x_1, y_1}^{x_2, y_2} P(x, y)\,dx + Q(x, y)\,dy. \tag{10}$$

The line integral (10) may be expressed as a simple integral by expressing y and dy in terms of x and dx by the equation of the curve, or we

may express all in terms of y and dy, or we may treat the two terms separately, one in terms of x and the other in terms of y. If the equation of the curve is in parametric form it may be convenient to express all in terms of the parameter and its differential. As an example let us take

$$\int_{2,\,0}^{6,\,4} y^2\,dx + x^2 dy$$

along the parabola treated above. In terms of x and dx we have

$$\int_2^6 [4x - 8 + x^2(x-2)^{-\frac{1}{2}}]\,dx,$$

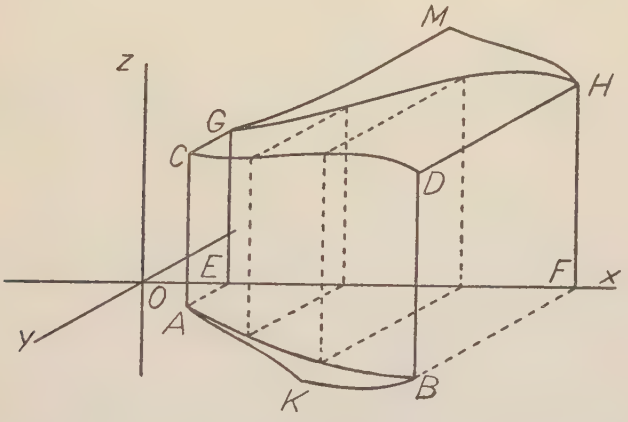


Fig. 108

and in terms of y and dy,

$$\int_0^4 \left[\frac{1}{2}y^3 + \frac{1}{16}(y^2+8)^2\right] dy.$$

The parabola is given by $x = 2 + t^2$, $y = 2t$ and in terms of t and dt the integral becomes

$$\int_0^2 [8t^3 + 2(2+t^2)^2]\,dt.$$

Some helpful notions involving the line integral are treated in the following paragraphs.

(a) **Geometrical interpretation.** Let x, y and z be the usual rectangular coordinates in three-dimensional geometry. The equation $y = f(x)$ represents a curve AB in the xy-plane (Fig. 108) or the

cylindrical surface ABDC, if interpreted in three dimensions. Let the surface $z = P(x, y)$ intersect this cylindrical surface in the space curve CD, then elimination of y between the equations $y = f(x)$ and $z = P(x, y)$ gives the equation of the cylindrical surface CDHG in the form $z = \phi(x)$. It is evident, then, that the line integral

$$\int_{AB} P(x, y)\, dx$$

taken along the curve $y = f(x)$ is equivalent to the simple integral

$$\int_{EF} \phi(x),\, dx,$$

or, in fact, to the plane area EFHG, which is the projection on the xz-plane of the cylindrical area ABDC. Of course the result will be positive or negative according to the order of the integration, that is,

$$\int_{AB} P(x, y)\, dx = -\int_{BA} P(x, y)\, dx.$$

The line integral from A to B will, in general depend on the curve followed. Corresponding to any other curve as AKB, we should have a curve GMH, and the integral

$$\int_{AKB} P(x, y)\, dx$$

is equal to the area EF HMG. Clearly, then, the area of the closed curve GHM is given by

$$\int_{AKB} P(x, y)\, dx + \int_{BA} P(x, y)\, dx = \int_{AKBA} P(x, y)\, dx.$$

Evidently, integration around any closed curve in the xy-plane is the area of a certain closed curve in the xz-plane, where the two closed curves are projections of the same closed curve drawn on the surface $z = P(x, y)$.

In a similar manner the integral

$$\int Q(x, y)\, dy$$

may be interpreted by area in the yz-plane, associated with a curve in the yz-plane which depends on the path of integration in the

xy-plane, both being projections of the same curve on the surface $z = Q(x, y)$.

In particular let the surface $z = P(x, y)$ be the plane $z = y$, then the curve in the xz-plane has the same area as that in the xy-plane, and is given by either form

$$\int z\,dx, \int y\,dx$$

taken around the boundary of the area. This is the familiar integral for area in the plane, but in the conventional case for plane geometry we should have

$$A = -\int y\,dx.$$

if the integral is taken around the boundary of the curve in a counter clockwise sense. Similarly,

$$A = \int x\,dy,$$

and if we add these two expressions for area of a closed curve we have

$$A = \frac{1}{2}\int x\,dy - y\,dx. \tag{11}$$

This is a line integral giving the area of a closed curve in the xy-plane, the integral to be taken around the boundary of the area in counter clockwise sense.

(b) **Green's Theorem.** Consider the iterated integral

$$\int_{x_1}^{x_2}\int_{f_1(x)}^{f_2(x)} \frac{\partial}{\partial y} P(x, y)\,dy\,dx. \tag{12}$$

Since the integrand is the partial derivative with respect to y of the function $P(x, y)$ we have after integrating with respect to y

$$\int_{x_1}^{x_2} P\,[x, f_2(x)]\,dx - \int_{x_1}^{x_2} P\,[x, f_1(x)]\,dx.$$

But this is equal to the sum of the line integrals

$$\int_{x_1}^{x_2} P(x, y)\,dx, \int_{x_2}^{x_1} P(x, y)\,dx,$$

where the first is taken along $y = f_2(x)$ from x_1 to x_2 and the second along $y = f_1(x)$ from x_2 to x_1. If the two paths form a closed curve, which we may denote by C, then the sum is the line integral of $P(x, y)$ taken in a clockwise sense around the curve, while (12) represents the double integral of $\frac{\partial P}{\partial y}$ over the area inclosed by the curve.

$$\iint_{(C)} \frac{\partial P}{dy}\, dy\, dx = -\int_{(C)} P\, dx, \tag{13}$$

where the line integral on the right is taken around the curve in a counter clockwise sense. Similarly it may be shown that for a function $Q(x, y)$ we have

$$\iint_{(C)} \frac{\partial Q}{\partial x}\, dx\, dy = \int_{(C)} Q\, dy. \tag{14}$$

Here the line integral on the right is equal to the double integral, as written, if taken around the curve in a counter clockwise sense. From equations (13) and (14) we have

$$\int_{(C)} P\, dx + Q\, dy = \iint_{(C)} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy. \tag{15}$$

This is known as **Green's theorem.** Referring to the geometrical interpretation in (a) above the closed curve C in the xy-plane is the trace of a cylindrical surface, and the line integral is the sum of two projected areas, the projection on the xz-plane of the area cut from $z = P(x, y)$ by the cylinder and the projection on the yz-plane of the area cut from $z = Q(x, y)$ by the cylinder. The double integral represents the volume bounded by the cylinder and the surfaces

$$z = 0 \text{ and } z = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

(b) **Exact differentials.** Consider a function of two variables, say

$$z = F(x, y). \tag{16}$$

The total differential of this function is

$$dz = \frac{\partial F}{\partial x}\, dx + \frac{\partial F}{\partial y}\, dy. \tag{17}$$

If we select a curve $y = f(x)$ in the xy-plane and integrate along this curve from (x_1, y_1) to (x_2, y_2) we have

$$\int_{z_1}^{z_2} dz = \int_{x_1, y_1}^{x_2, y_2} \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy. \tag{18}$$

The quantity on the right is a line integral and may be evaluated as indicated for the general case. But on the other hand, this quantity must represent merely the change in z in going from (x_1, y_1, z_1) to (x_2, y_2, z_2) on the surface (16). Therefore

$$\int_{z_1}^{z_2} dz = z_2 - z_1 = F(x_2, y_2) - F(x_1, y_1). \tag{19}$$

We may find the value of the line integral (18) by use of (19) when the function $F(x, y)$ is known. Given $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ we may write

$$F(x, y) = \int \frac{\partial F}{\partial x} dx + \phi(y) \tag{20}$$

where in integrating with respect to x we regard y as constant. Also

$$F(x, y) = \int \frac{\partial F}{\partial y} dy + \theta(x) \tag{21}$$

where x is regarded as constant when integrating with respect to y. The arbitrary functions $\phi(y)$ and $\theta(x)$ may be chosen so that (20) and (21) agree.

Now equation (19) shows that the line integral (18) depends only on the surface (16) and the two points (x_1, y_1) and (x_2, y_2), and that its value is not dependent on the path taken between these points. This is equivalent to the statement that the line integral

$$\int_{(C)} \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy \tag{22}$$

taken around a closed curve C must have the value zero. In fact, if we apply (15), putting $P = \frac{\partial F}{\partial x}$, $Q \frac{\partial F}{\partial y}$, then we have

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial^2 F}{\partial x \, \partial y} - \frac{\partial^2 F}{\partial y \, \partial x},$$

and the right member here is identically zero, so that, looking at the double integral in (15) we see that (22) must be zero.

We see then, that if $Pdx + Qdy$ is the total differential of some function $F(x, y)$ then

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0, \tag{23}$$

and the line integral (10) is independent of the path, depending only on the end points, and the line integral (15) around any closed curve is zero. Conversely, if (23) is fulfilled, then $Pdx + Qdy$ is the differential of some function $F(x, y)$ and the other conditions follow. Equation (23) is known as the **test for an exact differential,** and when it is fulfilled $Pdx + Qdy$ is said to be an **exact differential.**

It is very important to note that for an exact differential the line integral from a fixed point (x_1, y_1) along any path is a definite function of x and y, while in the general case it is not a function of x and y only, but depends also on the path of integration.

(d) **Extension to three variables.** The notion of line integral and the ideas associated with it are easily extended to cases involving three or more independent variables. We may state here some of the relations, confining attention to the case of three independent variables. If P, Q and R are functions of x, y and z then

$$\int_{AB} P\,dx + Q\,dy + R\,dz \tag{24}$$

may indicate a line integral taken along a curve connecting the two points, A and B, in space. Such a curve may be given by two equations of the form

$$y = f_1(x), \quad z = f_2(x)$$

and we may evaluate (24) as a simple integral by expressing y, dy, z and dz in terms of x and dx. If the curve is given in parametric form, all may be expressed in terms of the parameter and its differential.

If there is a function $w = F(x, y, z)$ such that

$$P = \frac{\partial F}{\partial x}, \quad Q = \frac{\partial F}{\partial y}, \quad R = \frac{\partial F}{\partial x},$$

then $Pdx + Qdy + Rdz$ is the total differential of w, and (24) will represent the change in value of w in going from A to B, and will be independent of the path, depending only on the coordinates of these two points. Necessary and sufficient conditions that there shall exist such a function w are:

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} = \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} = 0. \tag{25}$$

Equations (25) are the **conditions for an exact differential** of a function of three variables. If these equations are not fulfilled the line integral (24) will depend on the path as well as on the coordinates of the end points.

(e) **Physical applications.** (i) **Perfect gas.** For any gas there is a characteristic equation connecting the volume, pressure and absolute temperature of a unit quantity of the gas. Confining attention to the case of a perfect gas we may take

$$pv = BT \tag{26}$$

where B is a constant. The quantity of heat received by the gas when a change occurs in the values of the variables may be written

$$_1q_2 = \int_{p_1v_1T_1}^{p_2v_2T_2} c_v dT + Apdv. \tag{27}$$

Here c_v is "specific heat at constant volume," and is constant for a given perfect gas. The integral of $c_v dT$ represents the heat involved in the rise of temperature, while the integral of Apdv represents the mechanical work done by the gas in expanding. The constant A is to convert from units of mechanical work to heat units. The expression (27) is a line integral, and may be expressed in terms of any two of the variables by means of (26). If we choose p and v as the independent variables we have

$$_1q_2 = \int_{p_1,\, v_1}^{p_2,\, v_2} \frac{C_v}{B} vdp + \frac{C_v}{B} kpdv \tag{28}$$

where $k = \dfrac{c_v + AB}{c_v} = \dfrac{c_p}{c_v}$, the quantity c_p so defined being "specific heat at constant pressure," also constant.

Since k is greater than unity the quantity to be integrated in (28) is not an exact differential, and the quantity of heat q cannot be written as a function of the variables p and v. If pressure, volume and temperature vary through any series of values, then return to their initial values through a different series of values, the quantity q does not return to its initial value, but differs from its initial value by an amount depending on the character of the intermediate changes in pressure, volume and temperature. In other words, the line integral (28) depends on the path of integration.

Equation (27) may be written

$$_1q_2 = \int_{T_1}^{T_2} c_v dT + \int_{p_1, v_1}^{p_2, v_2} A p dv.$$

We may write $_1q_2 = {}_1q_2' + {}_1q_2''$, where

$$_1q_2' = \int_{T_1}^{T_2} c_v dT = c_v(T_2 - T_1), \tag{29}$$

$$_1q_2'' = \int_{p_1, v_1}^{p_2, v_2} A p dv. \tag{30}$$

The integral in (29) does not depend on the path of integration, and may be evaluated as shown. We may write $q' = c_v T + q_o'$, where q_o' is constant. If q' is divided by A, to convert to conventional units for energy, the result is the **intrinsic energy** of the gas. On the other hand, q'' cannot be written as a function of p and v, since (30) is a line integral which becomes definite only when p is written as a function of v, i.e., when the path is specified.

If the quantity to be integrated in (28) is divided by $T = \frac{1}{B} pv$, the resulting integral may be written

$$_1S_2 = \int_{p_1 v_1}^{p_2 v_2} C_v \frac{dp}{p} + C_v k \frac{dv}{v}. \tag{31}$$

The quantity S, called **entropy,** is a function of p and v, since the quantity to be integrated is an exact differential. We have

$$S = C_v \log_e (pv^k) + S_o, \tag{32}$$

where S_o is a constant.

The line integral (28) is to be taken along some path given by an equation of the form $f(p, v) = 0$. First let us consider the case of **isothermal expansion,** where the temperature remains constant while p and v vary according to the law $pv = C$. If we express the integral in terms of v and dv we have

$$_1q_2 = \int_{v_1}^{v_2} \frac{C_v}{B} (k - 1) C \frac{dv}{v} = \frac{C_v}{B} (k - 1) C \log_e \frac{v_1}{v_o} = AC \log_e \frac{v_1}{v_o}. \tag{33}$$

This is the quantity of heat necessary to maintain a constant temperature in the gas during the expansion from v_o to v_1.

Next, let us inquire along what path the change of heat is zero. We must have

$$vdp + kpdv = 0,$$

which gives

$$\frac{dp}{p} + k\frac{dv}{v} = 0$$

and integrating both expressions we have

$$\log p + k \log v = \log C_1$$

where $\log C_1$ is a convenient form for the constant of integration. The relation may be written

$$pv^k = C_1. \tag{34}$$

This is the equation of a path along which the line integral (28) or (27), is zero. When expansion takes place according to this law we see by equation (32) that the entropy remains constant, while the temperature changes, since (26) and (34) give

$$T = \frac{C_1}{B} v^{1-k}. \tag{35}$$

No heat being received or given off, the expansion is said to be **adiabatic.**

(ii) **Force field.** If an object moves under the action of a force, and if the three rectangular components of force when the object is at the point (x, y, z) are expressed as three functions X, Y and Z, each a function of x, y and z, then the work done upon the body by the variable force is the line integral

$${}_1W_2 = \int_{x_1\, y_1\, z_1}^{x_2\, y_2\, z_2} X\,dx + Y\,dy + Z\,dz. \tag{36}$$

If the condition (25) is fulfilled, then there exists a function W(x, y, z) such that

$$\frac{\partial W}{\partial x} = X, \quad \frac{\partial W}{\partial y} = Y, \quad \frac{\partial W}{\partial z} = Z,$$

and the line integral (36) represents changes in value of this function in going from one point to the other, and is independent of the path.

The function $U = -W$ is called the **potential** function, and the change of potential is the negative of the work done. Where such a function exists the force field is said to be a **conservative** field.

If the condition (25) is not fulfilled the line integral (36) will depend on the path, no potential function exists, and the field is **non-conservative.**

The notions involved here are readily applied to the case of a plane force field, dealing with two independent variables.

EXERCISE LIV

1. Evaluate $\int_0^6 \int_0^x \int_x^y (x^2 + y^2 + z^2)\, dz\, dy\, dx$, and give a geometrical interpretation.

2. Find the moment of inertia of a homogeneous cube with respect to an edge.

3. Find the center of gravity of a cube if the density varies as the distance from a certain edge.

4. Find the center of gravity of a sphere if the upper half is homogeneous and the density of the lower half varies as the distance from the center, the two parts being of equal weight.

5. Find the center of gravity and the radius of gyration about a diameter of the base of a right circular cylinder in which the density varies uniformly from w pounds per unit volume at the upper base to 2 w pounds per unit volume at the lower base. Express results in terms of the radius r and the altitude h.

6. Find the value of $\int\int\int xy\, d\, V$, extending the integral over the volume of a rectangular solid of dimensions a, b and c, using three concurrent edges as axes of reference.

7. A surface is given by the equation $xyz = 5$. Write an integral which expresses the area of that portion of this surface within the cylinder $(x - 2)^2 + (y - 3)^2 = 1$.

8. Find the mean value of the product xyz on that part of the surface of a unit sphere included in the first octant.

9. Find the volume and surface area of the part of cylinder $x^2 + y^2 = a^2$ intercepted by the cylinder $x^2 + z^2 = a^2$.

10. Which of the following are exact differentials? Find the functions whose total differentials appear.

(a) $x\, dx + y\, dy$;
(b) $y\, dx - x\, dy$;
(c) $(x - y)\, dx - x\, dy$;
(d) $x^2 dx - x\, y\, dy$
(e) $3x^2 y dx + (x^3 - y^3)\, dy$.
(f) $z\, dx + z\, dy + (x + y)\, dz$.

11. Find the value of the line integral

$$\int_1^3 (x^2 + y^2)dx,$$

taken along the path $y = x$.

12. Find the values of

$$\int_{o,\,o}^{1,\,1} y\,dx - x\,dy$$

along the three different paths $y = x$, $y = x^2$, $y = x^3$.

13. By an application of equation (11) use the results of example 12 to find areas inclosed by the paths.

14. By an application of Green's theorem, equation (15), show that the expression (11) is equal to the double integral for the area inclosed by the path.

15. Verify Green's theorem for the integral

$$\int y^2\,dx + x^2\,dy$$

taken around the rectangle whose vertices are (o, o), (2, o), (2, 3) and (o, 3).

16. In example 15 substitute the circle $x^2 + y^2 = 1$ for the rectangle.

17. Apply equation (28) to find the amount of heat received or given off when v changes from v_o to $2v_o$ along the path $pv^2 = C$. Express the result in terms of v_o, C and the constants of the equation.

18. What is the temperature change involved in example 17?

19. What change of entropy is involved in example 17?

20. Find the amount of heat transferred when the volume of the gas changes from v_1 to v_2 along the path $pv = C$, then from v_2 to v_1, along the path $pv^2 = C_1$, then back to the initial state along the path $v = v_1$, by the change of p from $\frac{C_1}{v_1^2}$ to $\frac{C}{v_1}$.

21. Discuss the change of entropy during an isothermal expansion of a perfect gas.

22. In a plane force field the two components of force on a unit mass are

$$X = -2x, \quad Y = -2y.$$

Show that the force field is conservative and find the potential function for this field, assuming that the potential is zero at the origin.

23. If the two components of force on a unit mass in a plane force field are

$$X = ay, \quad Y = bx,$$

show that the field is non-conservative unless $a = b$. Find the work done by the force on the unit mass in going around the rectangle whose sides are $x + y = 1, x - y = -1, x + y = -1$ and $x - y = 1$. Show that this reduces to zero if $a = b$. Find the potential function for the case when $a = b$. Show that no work is done along any path of the form $x^a y^b = C$, and show that when $a = b$ the potential function is constant along this path. Also show that the force at any point on such a path is perpendicular to the path.

24. In a certain three–dimensional force field the potential function is

$$\frac{-m}{\sqrt{x^2 + y^2 + z^2}}$$

show that the force varies inversely as the square of the distance from the origin and is directed toward the origin.

BOOKS FOR SUPPLEMENTARY READING

Solid Geometry, Charles Smith (Macmillan, 1903).
Advanced Calculus, W. F. Osgood (Macmillan, 1925), Chapters IV, V, VI, VIII, XI, XII, and XVIII.
Advanced Calculus, E. B. Wilson (Ginn, 1912), Chapters IV, V and XII.
Mathematical Analysis, Goursat-Hedrick, Vol. I (Ginn, 1904), Chapters VII, X, XI, and XII.
Integral Calculus, W. E. Byerly (Ginn, 1888).

CHAPTER XIII

TREATMENT OF EMPIRICAL DATA

Art. 58. Theory of Probability

Quantities that are assigned values as a result of observations or measurements are not usually known exactly, but only estimated approximately. For example, the temperature at which a solution of salt in water of given concentration will freeze is not definitely known. If one man makes a single determination, his result is subject to error from several causes, such as incorrect calibration of the thermometer, inaccuracy in reading the thermometer, inaccuracy in weighing salt and water, loss of water by evaporation, etc. From a single determination there is no way of estimating the amount of error, but if two or more determinations are made independently, the closeness with which these may agree permits of an estimate of the reliability of results. If two independent determinations are made, the two will probably not agree, but we have an instinctive feeling that the true value is more likely to lie between the two values obtained than it is to lie elsewhere, and there is a feeling that the accuracy or precision of the determinations is in some way associated with the difference between the two results.

When a series of observations have been made to determine the value of a certain unknown quantity, it is of vital importance to know not only the most probable value of the quantity, but also a measure of precision of the result. It becomes necessary to have a definite technical definition of probability and for this purpose we begin with simple ideal cases.

1. **Simple ideal cases.** Suppose that of n different events one must occur, only one can occur and all are equally likely to occur, then if r of these are favorable events and the remaining $n - r$ events are unfavorable, the **probability** of the event which occurs being favorable is $\frac{r}{n}$ and the **probability** of its being unfavorable is $\frac{n - r}{n}$. For example, if a coin is tossed it must fall heads or tails, not both, and the two events are equally likely to occur. The probability of heads is $\frac{1}{2}$, that of tails is $\frac{1}{2}$. If two coins are tossed, they may fall both heads,

first head and second tail, first tail and second head, or both tails. Each of these four events are equally likely. The probability of both heads is $\frac{1}{4}$, of both tails is $\frac{1}{4}$, but that of one head and one tail is $\frac{2}{4}$ or $\frac{1}{2}$. If an event is certain to occur, its probability is unity.

Two events are said to be **independent** if the probability of one does not depend on whether the other occurs or fails to occur. A **compound** event is the joint occurrence of two or more events. Several cases are of importance.

(a) **Concurring independent events.** The probability that several independent events will all occur is the product of their separate probabilities.

Consider two events, and let it be possible for the first to happen in m_1 ways and fail in n_1 ways, all equally likely, then its probability is $p_1 = \frac{m_1}{m_1 + n_1}$, and with similar meaning for the notation let the probability of the second event be $p_2 = \frac{m_2}{m_2 + n_2}$. Now with each of the m_1 ways that the first may happen, the second may happen in m_2 ways, so that the two may happen together in $m_1 m_2$ ways. But the first event may be disposed of in $m_1 + n_1$ ways and the second in $m_2 + n_2$ ways, so that the total number of ways of disposing of the two events is $(m_1 + n_1)(m_2 + n_2)$. The probability of the two events happening together is therefore

$$\frac{m_1 m_2}{(m_1 + n_1)(m_2 + n_2)} = p_1 p_2.$$

(b) **Mutually exclusive events.** If only one of several events can occur, the probability that one will occur is the sum of their separate probabilities.

Consider two events, and suppose that the first may happen and the other fail in m_1 ways, that the second may happen and the first fail in m_2 ways, and that both may fail in m_3 ways. Then the probability that the first shall happen is $p_1 = \frac{m_1}{m_1 + m_2 + m_3}$, that the second shall happen, $p_2 = \frac{m_2}{m_1 + m_2 + m_3}$, that the first or second shall happen $p = \frac{m_1 + m_2}{m_1 + m_2 + m_3} = p_1 + p_2$.

To illustrate these two propositions consider a pack of 52 playing

cards. If one card is drawn at random the probability that it is a spade is $\frac{1}{4}$, the probability that it is a king is $\frac{1}{13}$, the probability that it is the king of spades is $\frac{1}{4} \times \frac{1}{13} = \frac{1}{52}$. Similarly, the probability that it is the king of diamonds is $\frac{1}{52}$, of hearts $\frac{1}{52}$, of clubs $\frac{1}{52}$ and the probability that it is a king is $\frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{1}{13}$. In drawing four cards, the probability that all are kings is the number of ways to draw four kings divided by the total number of ways to draw four cards. Now there are only four kings, but they may be drawn in 4! different orders, and the number of ways to draw four from 52, allowing for different orders, is $52 \cdot 51 \cdot 50 \cdot 49$, therefore the probability of drawing four kings is $\frac{4 \cdot 3 \cdot 2 \cdot 1}{52 \cdot 51 \cdot 50 \cdot 49}$. The probability of drawing four spades is $\frac{13 \cdot 12 \cdot 11 \cdot 10}{52 \cdot 51 \cdot 50 \cdot 49}$. The probability of drawing four from one suit is $\frac{4 \cdot 13 \cdot 12 \cdot 11 \cdot 10}{52 \cdot 51 \cdot 50 \cdot 49}$.

(c) **Repeated trials.** If the probability of success at each trial is p, and the probability of failure is q, then the probability of succeeding exactly r times in n trials is the term involving p^r in the binomial expansion of $(p + q)^n$, that is,

$$\frac{n!}{r!(n - r)!} p^r q^{n-r}.$$

To see the truth of this we note that the r trials that are to be successful may be selected from the n trials in $\frac{n!}{r!(n - r)!}$ different ways, and for each way, the probability that the r selected shall all be successful is the product of their separate probabilities, p^r, and the probability that the remaining $n - r$ trials will fail is q^{n-r}.

To illustrate, the probability of throwing an ace in a single trial with one die is $\frac{1}{6}$, and if four trials are made we have

$$\left(\frac{1}{6} + \frac{5}{6}\right)^4 = \left(\frac{1}{6}\right)^4 + 4\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right) + 6\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 + 4\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3 + \left(\frac{5}{6}\right)^4.$$

These five terms are respectively the probabilities of 4 aces, 3 aces, 2 aces, one ace and no ace in four throws. The sum of these five

probabilities is unity, since one and only one of these five events must happen. The same expressions also represent the probabilities if the trials are made simultaneously by throwing four dice at once.

Again, suppose five coins are tossed. The probability that each will fall head is $\frac{1}{2}$. We write then

$$\left(\frac{1}{2}+\frac{1}{2}\right)^5 = \left(\frac{1}{2}\right)^5 + 5\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right) + 10\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 + 10\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 + 5\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^5$$

$$= \frac{1}{32} + \frac{5}{32} + \frac{10}{32} + \frac{10}{32} + \frac{5}{32} + \frac{1}{32}.$$

These six terms are respectively the probabilities of 5, 4, 3, 2, 1 and no heads among the five coins.

2. **Probability of errors of observation.** No rigorous demonstration is practicable for the law of probability as applied to errors of observation. The number of possibilities of various events, causes, etc., cannot be counted, so that the ideal cases discussed above do not apply. However, a well accepted theory of probability is in general use, and is found to give valuable aid in the treatment of empirical data. This theory may be shown to be consistent with the elementary conception of probability introduced above.

An observation or measurement is subject to error due to many causes, and each of these causes may be further analyzed into other causes, and so on, indefinitely. Let us suppose that the sources of error are all capable of being evaluated as multiples of a certain elementary source which produces an error of magnitude a. Let us suppose that the total number of such elementary sources involved is n, and that each is as likely to produce a negative error as a positive one. Then if the number of these which fall positive exceeds the number which fall negative by m, we shall have a positive error of magnitude ma. The situation is similar to that when n coins are tossed. We write

$$\left(\frac{1}{2}+\frac{1}{2}\right)^n = \frac{1}{2^n} + n\frac{1}{2^n} + \frac{n(n-1)}{1\cdot 2}\frac{1}{2^n} + \cdots + \frac{n(n-1)\cdots(n-r+1)}{1\cdot 2\cdots r}\frac{1}{2^n} + \cdots + \frac{1}{2^n}.$$

The first term is the probability that all elementary sources of error will have negative effects, i.e., that the error will be $-na$. The term $\frac{n(n-1)\ldots(n-r+1)}{1\cdot 2\cdot 3\ldots} \frac{1}{2^n}$ is the probability that r will be negative, and $n - r$ positive, or that the error will be $(n - 2r)a$.

For the case when $n = 5$ these numbers are the same as those for the case when five coins are tossed. We may make a graph of the

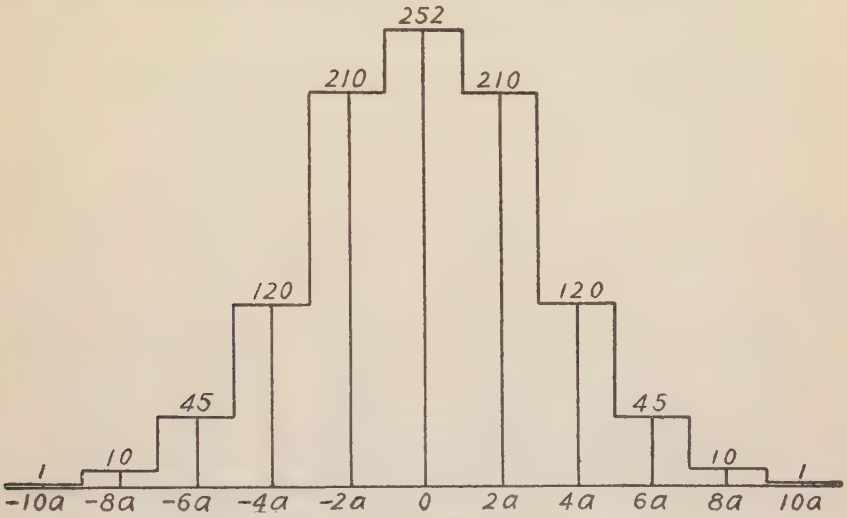


FIG. 109

probability function for the case when $n = 10$. If p denotes the probability of an error x, then we have the following table of values:

x:	$-10a$	$-8a$	$-6a$	$-4a$	$-2a$	0	$2a$
p:	$\frac{1}{1024}$	$\frac{10}{1024}$	$\frac{45}{1024}$	$\frac{120}{1024}$	$\frac{210}{1024}$	$\frac{252}{1024}$	$\frac{210}{1024}$

	$4a$	$6a$	$8a$	$10a$
	$\frac{120}{1024}$	$\frac{45}{1024}$	$\frac{10}{1024}$	$\frac{1}{1024}$

The sum of these eleven values of p is unity. In Figure 109 ordinates are drawn, at each of these values of x, proportional to the corresponding values of p. If rectangles are drawn of width $2a$ and if the altitudes are taken as $y = \frac{p}{2a}$, then the area of each rectangle is $p = 2ay$, the

probability that the error will fall on its base, and the total area is unity.

Now in an actual observation, the source of error does not consist of 10 equal elementary sources, nor of any finite number of equal sources, but as the number n above is indefinitely increased, the hypothetical situation above becomes more and more nearly a correct representation of the actual case. We wish to find a function, $y = f(x)$, such that the differential area $y\,dx$ is the probability that the error will fall on the differential base dx. To do this we start with a finite n, take Δx as 2a, and finding the corresponding Δy, we find the limit of $\frac{\Delta y}{\Delta x}$ as n

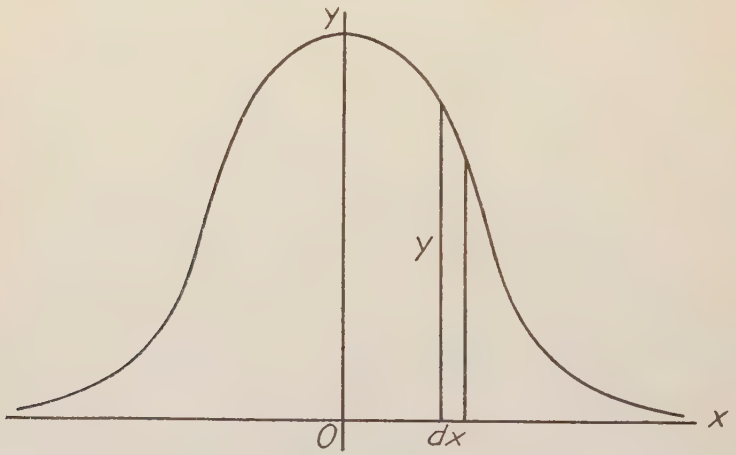


FIG. 110

increases indefinitely, and from this we obtain y as a function of x (Fig. 110).

Two successive values of x, differing by 2a, are $x_1 = (n - 2r)a$ $x_2 = (n - 2r - 2)a$, and the corresponding values of y are

$$y_1 = \frac{n(n-1)\ldots(n-r+1)}{(2a)\,1\cdot 2\cdot 3\ldots r}\,\frac{1}{2^n},\quad y_2 = \frac{n(n-1)\ldots(n-r)}{(2a)\,1\cdot 2\cdot 3\ldots r+1}\,\frac{1}{2^n}.$$

By division we have

$$\frac{x_1}{x_2} = \frac{n-2r}{n-2r-2},\quad \frac{x_2 - x_1}{x_1} = \frac{-2}{n-2r}.$$

Similarly we have

$$\frac{y_1}{y_2} = \frac{r+1}{n-r'} \quad \frac{y_2 - y_1}{y_1} = \frac{n-2r-1}{r+1}$$

Eliminating r we are able to write

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{-2x_1y_1 - y_1(x_2 - x_1)}{\frac{n}{2}(x_2 - x_1)^2 + x_1(x_2 - x_1) + (x_2 - x_1)^2}.$$

As n increases indefinitely, the difference quotient on the left approaches the derivative, $\frac{dy}{dx}$, and remembering that $x_2 - x_1 = 2a$, we have

$$\frac{dy}{dx} = \lim_{n \to \infty} \frac{-2xy - 2ay}{2na^2 + 2ax + 4a^2}.$$

Now a decreases as n increases, approaching zero as n becomes infinite, and the term $2na^2$ becomes indeterminate. We see that as the number n of elementary sources of error is increased, the magnitude a of the error due to each elementary source must decrease, while the product na, the greatest possible error, will increase and the probability of this greatest error, $\frac{1}{2^n}$, rapidly decreases. The limit of na^2 must be definite for each set of circumstances under which observations are made, but will be different for different circumstances. Let us write

$$\frac{1}{2h^2} = \lim_{n \to \infty} na^2,$$

where h is a constant to be determined for each case treated, then we have

$$\frac{dy}{dx} = -2h^2xy, \text{ or } \frac{dy}{y} = -2h^2x\,dx.$$

Integrating we have an equation which defines y as a function of x,

$$\log y = -h^2x^2 + c, \text{ or } y = ke^{-h^2x^2}, \qquad (1)$$

where $k = e^c$, is a convenient form for the constant of integration.

One condition that our curve must fulfill is that the total area under the curve must be unity. We see that the curve is symmetrical with

respect to the y-axis, and that the x-axis is an asymptote, therefore the area is

$$2\int_0^\infty ke^{-h^2x^2}\,dx = 1.$$

It is not possible to express in terms of elementary functions any function of x whose derivative is the integrand here, but by expanding the integrand in a power series and integrating term by term we may obtain values, as accurate as desired, for any value of the upper limit. The definite integral written above is given in most tables of integrals.

$$2\int_0^\infty ke^{-h^2x^2}\,dx = \frac{k\sqrt{\pi}}{h} = 1,$$

therefore $k = \frac{h}{\sqrt{\pi}}$. The probability that an error will lie between $-x$ and x, i.e., that it will be not greater numerically than x, is

$$P = 2\int_0^x y\,dx = \frac{2h}{\sqrt{\pi}}\int_0^x e^{-h^2x^2}\,dx = \frac{2}{\sqrt{\pi}}\int_0^{hx} e^{-(hx)^2}\,d(hx). \quad (2)$$

The values of this integral are tabulated for various values of hx in many physical and mathematical tables. It is known as the **probability integral.**

This integral is, of course, a function of its upper limit the product hx. The probability that an error shall lie within certain limits, say from $-x$ to x, increases with h. The constant h, therefore, is a **measure of the precision** of the observation. If h is large the repeated observations are likely to lie close together and near the true value, while if h is small they are likely to be widely scattered and far from the true value.

EXERCISES LV

1. What is the probability of throwing two sixes in a single throw with two dice?

2. In drawing two cards from the 52, what is the probability that both will be hearts? Ans. $\frac{1}{17}$.

3. If six coins are tossed, what is the probability that three will fall heads? Ans. $\frac{5}{16}$.

4. If six coins are tossed, what is the probability at least three will fall heads? Ans. $\frac{21}{32}$.

5. If a bag contains 10 white balls and 6 black balls, and two are drawn at random, what is the probability that one will be white and the other black? Ans. $\frac{1}{2}$.

6. In example 5 what is the probability that both balls will be white? That both will be black? Ans. $\frac{3}{8}; \frac{1}{8}$.

7. If five people draw one ball each from the bag in example 5, each returning the ball before the next draws, what is the probability that all will draw black balls? Ans. $\frac{243}{32768}$.

8. If in example 7 the balls are not returned to the bag, what is the probability that all will be black? Two black and three white? Ans. $\frac{1}{728}; \frac{75}{182}$.

9. If a gunner has an average of three hits in 10 trials, what is the probability that he will make at least one hit in five trials? (Note that $p = 1 - q$, where q is the probability of failing in all five trials, i.e., $p = 1 - \left(\frac{7}{10}\right)^5$).

10. In example 9, what is the least number of trials such that the probability of making at least one hit is greater than that of failing in all trials? Ans. 2.

11. Plot the probability curve $y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2}$ for $h = 1$ and for $h = \frac{1}{2}$ on the same axis.

12. Supposing h to be known, find from the table of the probability integral a value of x, in terms of h, such that an error is as likely to be numerically less than x as it is to be numerically greater than x.

13. Find from the table how many observations out of 20 should lie within $\frac{3}{5h}$ units of the true value. Ans. 12.

Art. 59. Method of Least Squares

1. **Direct observations on a single quantity.** Suppose a series of n observations are made, all under similar conditions, to determine the value of a certain unknown quantity Q. The theoretical probability

of a certain error x is of course zero, according to the theory of probability, since

$$\frac{h}{\sqrt{\pi}} \int_{x_1}^{x_2} e^{-h^2x^2}\, dx$$

represents the probability that an error will fall between x_1 and x_2. The probability pertains to an interval, then, and is the area above the interval. But practically, there is always a smallest division to which readings or observations are made, and if this is denoted by c, the probability of a certain error x is really the probability that it will be on an interval of length c. For example, if we are reading a scale to tenths, we read .8 if this seems nearer than .7 or .9; that is, .8 means between .75 and .85. The probability, then of a certain error x is

$$p_x = \frac{ch}{\sqrt{\pi}} e^{-h^2x^2}.$$

Now suppose the n observations referred to above give the values $m_1, m_2, \ldots . m_n$ for the quantity Q. We wish to find a value M which is the most probable value of Q. If M is the value then $x_1 = M - m_1$, $x_2 = M - m_2 \ldots .$, $x_n = M - m_n$ are the errors and their probabilities are

$$\frac{c\,h}{\sqrt{\pi}} e^{-h^2x_1^2}, \ldots . , \frac{c\,h}{\sqrt{\pi}} e^{-h^2x_n^2},$$

and the probability that this particular set of errors should occur is the product of their separate probabilities,

$$P = \frac{c^n h^n}{\pi^{\frac{n}{2}}} e^{-h^2(x_1^2 + x_2^2 + \ldots . + x_n^2)}.$$

It is clear that the most likely value M of Q is that which makes the set of errors most probable, and therefore that which makes the quantity $x_1^2 + x_2^2 + \ldots . + x_n^2$ a minimum. Therefore M is to be chosen so that

$$(M - m_1)^2 + (M - m_2)^2 + \ldots . + (M - m_n)^2$$

is a minimum. Taking the derivative with respect to M and equating to zero we find

$$M = \frac{m_1 + m_2 + \ldots . + m_n}{n}. \qquad (1)$$

Thus the arithmetical mean of the n measurements is the most probable value of Q as determined by this series of observations, but it is not the true value of Q, and the quantities $M - m_1$, $M - m_2$, etc., are not called the errors, but the **residual errors,** or **residuals,** and are denoted by $v_1, v_2, \ldots . v_n$. The principle of **Least Squares,** then is that M is so chosen that

$$\Sigma v^2 = v_1^2 + v_2^2 + \ldots . + v_n^2$$

is a minimum.

Having found the most probable value of the unknown quantity Q, we now wish to determine an estimate of the precision of this result. We saw in the last article that the quantity h is a measure of precision and it is now necessary to find a value in terms of the residuals, v_1, v_2, etc., of the quantity h.

Our information consists merely of the n measurements, $m_1, m_2, \ldots . m_n$ and we wish to determine h so that the area under the curve

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2}$$

shall be distributed horizontally so as to conform to the distribution of these measurements. There being n measurements, and the total area being unity, the area assigned to each measurement is $\frac{1}{n}$. The origin for x is Q, and in selecting M, our estimate of Q, we have fixed the origin of the v's, our best estimates of the x's, so that the first moment of the total area about the vertical axis through M is zero. That is,

$$\frac{v_1}{n} + \frac{v_2}{n} + \ldots . + \frac{v_n}{n} = 0.$$

Now we may select h so that the second moment, or moment of inertia, of the theoretical area about the y-axis is equal to that of the actual area found. If M differs from Q by q, then we should have

$$\frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 e^{-h^2x^2} dx = \frac{\Sigma v^2}{n} + q^2,$$

and by the integral tables this is found to give

$$\frac{1}{2h^2} = \frac{\Sigma v^2}{n} + q^2,$$

where q is a quantity still to be determined. Of course q cannot be determined absolutely, but we can find from a given set of observations a reasonable estimate for this error.

First we note that if r is so chosen that

$$\frac{2}{\sqrt{\pi}} \int_0^{hr} e^{-h^2x^2}\, d(x) = 1/2,$$

then an error x is as likely to be numerically greater than r as it is to be numerically less than r. This value r is called the **probable error of a single observation,** and is found from the table to be

$$r = \frac{0.47694}{h}. \tag{2}$$

Now consider again the n observations $m_1, m_2, \ldots, m_n$. For convenience, let us suppose that $n = 9$, then M is determined by the condition that

$$h^2(M - m_1)^2 + h^2(M - m_2)^2 + \ldots + h^2(M - m_9)^2$$

shall be a minimum, and this gives

$$M = \frac{m_1 + m_2 + \ldots + m_9}{9}.$$

Let $5M_1 = m_1 + m_2 + m_3 + m_4 + m_5$, then

$$M = \frac{5M_1 + m_6 + m_7 + m_8 + m_9}{9}$$

and this is the value of M which gives a minimum for

$$h_1^2(M - M_1)^2 + h^2(M - m_6)^2 + \ldots + h^2(M - m_9)^2$$

where $h_1^2 = 5h^2$. We see then that the mean M_1 of 5 observation when combined with single observations, must be assigned a measure of precision h_1 which is related to the measure of precision h by the equation $h_1 = \sqrt{5}h$. The probable error of a single observation being r, and that of M_1 being r_1, we have

$$r = \frac{0.47694}{h}, \quad r_1 = \frac{0.47694}{h_1}, \quad r_1 = \frac{r}{\sqrt{5}}$$

Reasoning in this way, we find that the probable error of the mean of n observations is

$$r_o = \frac{r}{\sqrt{n}} = \frac{0.47694}{h\sqrt{n}}, \tag{3}$$

where r is the probable error of a single observation.

This quantity r_o is not necessarily the same as the quantity q considered above, but it is clear that q should vary with h and with n in the same manner as r_o. We will therefore assume q to have the form

$$q^2 = \frac{b^2}{h^2 n},$$

and consider further what is the most suitable value for b. We now have

$$\frac{1}{2h^2} = \frac{\Sigma v^2}{n} + \frac{b^2}{h^2 n},$$

whence we obtain

$$\frac{1}{h^2} = \frac{2\Sigma v^2}{n - 2b^2}.$$

Now b should be so chosen that this will give a reasonable value to h for all values of n. If $2b^2$ is different from unity, h becomes infinite when $n = 1$, and this would imply that $r = o$, which is not reasonable. If $2b^2 = 1$, h becomes indeterminate at $n = 1$, which is reasonable, since a single observation can furnish no criterion as to its precision. Therefore we assign b the value $\frac{1}{\sqrt{2}}$, which gives

$$q^2 = \frac{1}{2h^2 n} = \frac{1}{2(0.47694)^2}\, r_o^2.$$

We have, then, finally,

$$\frac{1}{h} = \sqrt{2}\sqrt{\frac{\Sigma v^2}{n - 1}}, \tag{4}$$

$$r = 0.6745\sqrt{\frac{\Sigma v^2}{(n - 1)}}, \tag{5}$$

$$r_o = 0.6745\sqrt{\frac{\Sigma v^2}{(n - 1)n}}. \tag{6}$$

These last two formulas, the probable error of a single observation and that of the mean of the set of observations, are the important working formulas in the discussion of precision. The factor $(\Sigma v^2)^{1/2}$ must be computed for each set of observations, but the remaining factors of both r and r_o are recorded in tabular form.

Consider the following example: A line was measured eight times with a tape twenty meters long, graduated to centimeters. The results of the eight measurements, in meters, are shown in the first column of the table below. The mean of these is found to be 188.875 meters, and the residuals are recorded in the second column, their squares in the third column. We find that the sum of the squares of

Observations	v	v^2
188.97	0.095	.009025
.88	.005	.000025
.91	.035	.001225
.99	.115	.013225
.83	−.045	.002025
.80	−.075	.005625
.81	−.065	.004225
.81	−.065	.004225
M = 188.875	Σv = .500	Σv^2 = .039600

the residuals is

$$\Sigma v^2 = 0.0396,$$

then by the aid of the tables we have

$$r = .2549\sqrt{0.0396} = 0.051,$$

$$r_o = .0901\sqrt{0.0396} = 0.018.$$

The result should now be recorded

$$M = 188.875 \pm 0.018 \text{ meters.}$$

The probable error is recorded to two digits, and the value of M is given to as many places as this reaches.

2. **Weighted observations of a single quantity.** Let us suppose that several independent determinations of a quantity Q have been made, with different degrees of precision. Let the results be m_1, m_2, m_3 . . . . m_n, and let measures of precision be respectively h_1, h_2, ... h_n. If M is the value assumed for Q, the errors are $x_1 = M - m_1$, $x_2 = M - m_2$, etc., and their probabilities may be denoted by

$$p_1 = \frac{c_1 h_1}{\sqrt{\pi}} e^{-h_1^2 x_1^2}, \; p_2 = \frac{c_2 h_2}{\sqrt{\pi}} e^{-h_2^2 x_2^2}, \text{ etc.}$$

and the probability of the set of errors is the product

$$p_1p_2 \ldots . p_n = \frac{c_1c_2 \ldots . c_nh_1 \ldots . h_n}{(\sqrt{\pi})^n} e^{-h_1^2x_1^2 + \ldots . + h_n^2x_n^2}.$$

Just as before, this will be a maximum when the negative exponent of e is a minimum, that is, when

$$h_1^2(M - m_1) + h_2^2(M - m_2) + \ldots . + h_n^2(M - m_n) = 0.$$

This gives for the most probable value of M the quantity

$$M = \frac{h_1^2m_1 + h_2^2m_2 + \ldots . h_n^2m_n}{h_2^2 + h_2^2 + \ldots . h_n^2}.$$

This quantity M is called the **weighted** mean, the results, m_1, m_2, etc., being given **weights,** or influences, proportional to the quantities h_1^2, h_2^2, etc. It is convenient to give these weights the notations w_1, w_2, etc., and write

$$\frac{w_1}{h_1^2} = \frac{w_2}{h_2^2} = \ldots . = \frac{w_n}{h_n^2} = \frac{1}{h^2}, \tag{7}$$

where h is the measure of precision of a determination of weight unity. We now have

$$M = \frac{w_1m_1 + w_2m_2 + \ldots . w_nm_n}{w_1 + w_2 + \ldots . w_n}. \tag{8}$$

It will be noticed that in assigning a weight w to a given determination, observation or measurement we say, in effect, that this determination is as reliable, i.e., has the same precision, as the mean of w determinations of weight unity. In practice the w's are taken as integers, approximately proportional to the squares of the h's, or inversely proportional to squares of the probable errors.

The probable error, r, of a determination of weight unity and the probable error, r_o, of the weighted mean M are now given by

$$r = 0.6745\sqrt{\frac{\Sigma wv^2}{n - 1}}, \tag{9}$$

$$r_o = 0.6745\sqrt{\frac{\Sigma wv^2}{(n - 1)\ \Sigma w}}, \tag{10}$$

where

$$\Sigma wv^2 = w_1(M - m_1)^2 + w_2(M - m_2)^2 + \ldots . w_n(M - m_n)^2,$$

$$\Sigma w = w_1 + w_2 + \ldots . w_n.$$

These summations must be computed for each set of data, but the remaining factor in r and r_o is the same as that which appeared before, and is found in the tables.

The formulas for r and r_o just given may be derived by reasoning similar to that used in 1, but let it suffice here to observe that these formulas are consistent with those previously derived when the w's are taken all equal, and reduce to the same when $w_1 = w_2 = \ldots . w_n = 1$. Moreover, the probable error, r, associated with weight unity, is increased if larger multiples are used as weights, while the use of larger or smaller multiples does not affect r_o.

3. **Linear functions of several quantities.** Let us suppose that several quantities, say for convenience three quantities, x, y and z, are to be determined, and that for this purpose several linear functions of these are measured, with the results

$$\begin{aligned} a_1x + b_1y + c_1z &= M_1, \\ a_2x + b_2y + c_2z &= M_2, \\ \ldots &\ldots \\ a_nx + b_ny + c_nz &= M_n. \end{aligned} \qquad (11)$$

Now these measurements M_1, M_2, etc., are subject to error, and, the number of equations being greater than the number of unknowns, it cannot be expected that values of x, y and z can be found to satisfy all equations, while it is likely that the most suitable values will not exactly satisfy any of the equations. Suppose x, y and z to have their most probable values, and let

$$\begin{aligned} a_1x + b_1y + c_1z - M_1 &= v_1, \\ a_2x + b_2y + c_2z - M_2 &= v_2, \\ \ldots &\ldots \\ a_nx + b_ny + c_nz - M_n &= v_n. \end{aligned} \qquad (12)$$

where v_1, v_2, etc., are the residuals, the quantities by which M_1, M_2, etc., differ from their most probable values. If h is the measure of precision of M_1 the probability of M_1 is

$$p_1 = \frac{ch}{\pi^{1/2}} e^{-h^2 v_1^2},$$

and if the n measurements are made with equal precision the probability of the set of measurements is

$$P = \frac{c^n h^n}{\pi^{\frac{n}{2}}} e^{-h^2(v_1^2 + v_2^2 + \ldots + v_n^2)}.$$

The most probable values of x, y and z are those which make this a maximum, and therefore $v_1^2 + v_2^2 + \ldots + v_n^2$ a minimum. A necessary condition is that the partial derivatives of this quantity with respect to x, y and z shall be zero, and these three equations are sufficient to determine x, y and z. To write these equations we observe that the partial derivative with respect to x is

$$2v_1 \frac{\partial v_1}{\partial x} + 2v_2 \frac{\partial v_2}{\partial x} + \ldots + 2v_n \frac{\partial v_n}{\partial x} = 2a_1 v_1 + 2a_2 v_2 + \ldots + 2a_n v_n$$

while similar forms hold for y and z, so that, dividing by 2 and substituting for v_1, v_2, etc., from equations (12), we have the **normal equations**

$$\begin{aligned} [aa]x + [ab]y + [ac]z &= [aM], \\ [ab]x + [bb]y + [bc]z &= [bM], \\ [ac]x + [bc]y + [cc]z &= [cM], \end{aligned} \qquad (13)$$

in which

$$\begin{aligned} [aa] &= a_1^2 + a_2^2 + \ldots + a_n^2, \\ [ab] &= a_1 b_1 + a_2 b_2 + \ldots + a_n b_n, \\ [ac] &= a_1 c_1 + a_2 c_2 + \ldots + a_n c_n, \\ [aM] &= a_1 M_1 + a_2 M_2 + \ldots + a_n M_n, \text{ etc.} \end{aligned}$$

In case the quantities M_1, M_2, etc., are determined with unequal precision, the equations (11) are assigned weights proportional to the

squares of the precision measures, and denoting these weights by w_1, w_2, etc., the quantity to be made a minimum is now

$$w_1v_1^2 + w_2v_2^2 + \ldots + w_nv_n^2.$$

The normal equations corresponding to (13) above now become

$$\begin{aligned} [waa]x + [wab]y + [wac]z &= [waM], \\ [wab]x + [wbb]y + [wbc]z &= [wbM], \\ [wac]x + [wbc]y + [wcc]z &= [wcM], \end{aligned} \tag{14}$$

where now

$$[waa] = w_1a_1^2 + w_2a_2^2 + \ldots + w_na_n^2,$$

$$[wab] = w_1a_1b_1 + w_2a_2b_2 + \ldots + w_na_nb_n, \text{ etc.}$$

It should be noted that, while we have referred to the precision of the observations or measurements in deriving normal equations, it is not necessary that the measures of precision be known in order that the normal equations may be used. Equations (13) do not require precision measures, and involve only the observation equations (11). The weights assigned in obtaining (14) may be merely estimates of the relative reliability of the observation equations, based on previous knowledge of the observer, of instruments used, of special conditions present, etc. In the absence of such basis for estimates, the form (13) should be used.

The method developed here for three unknowns, x, y and z, is obviously applicable to linear observation equations involving any number of unknown quantities, the number of normal equations of type (13) or (14), in each case, being equal to the number of quantities to be determined.

It is desirable to state here, without proof, a method for estimating the precision of quantities determined by normal equations. Let the values of x, y and z found from (14) be used in (12) to determine the values of the v's, then the probable error of an observation of weight unity is given by the formula

$$r = 0.6745\sqrt{\frac{\Sigma wv^2}{n-3}}.$$

If form (13) is used, all weights being assumed as unity, then r is the probable error of each of the measured quantities, M_1, M_2, etc. In

general, for n observations on q unknowns, the formula for probable error of an observation of weight unity is

$$r = 0.6745 \sqrt{\frac{\Sigma wv^2}{n - q}}. \tag{15}$$

If $q = 1$ this reduces to (9), as it evidently should.

Now suppose the solution of (14) to be written as linear functions of the quantities on the right, in the form

$$x = A_1[waM] + A_2[wbM] + A_3[wcM],$$

$$y = B_1[waM] + B_2[wbM] + B_3[wcM],$$

$$z = C_1[waM] + C_2[wbM] + C_3[wcM],$$

then the probable errors of x, y, and z are

$$r_x = \sqrt{A_1}r, \quad r_y = \sqrt{B_2}r, \quad r_z = \sqrt{C_3}r. \tag{16}$$

If equations (14) are solved by determinants in the usual way, and if the determinant of the coefficients of the unknowns is denoted by Δ, then evidently

$$A_1 = \frac{\Delta_a}{\Delta}, \quad B_2 = \frac{\Delta_b}{\Delta}, \quad C_3 = \frac{\Delta_c}{\Delta}, \tag{17}$$

where Δ_a, Δ_b, and Δ_c are the cofactors of [waa], [wbb] and [wcc], respectively, in the determinant Δ. The reciprocals of A_1, B_2 and C_3 are known as the weights of the respective unknowns. The manner of extending (16) and (17) to any number of unknowns is evident.

As an example, the angles x, y and z of a triangle were measured, x by four measurements with a mean value of 36° 25′ 47″, y by two measurements, mean 90° 36′ 28″, z by three measurements, mean 52° 57′ 57″. Since $z = 180 - (x + y)$, the observation equations are

$$x = 36° 25′ 47″, \text{ weight } 4,$$

$$y = 90° 36′ 28″, \text{ weight } 2,$$

$$x + y = 127° 2′ 3″, \text{ weight } 3.$$

The normal equations are found to be

$$7x + 3y = 526° 49′ 17″.$$

$$3x + 5y = 562° 19′ 5″.$$

and solving for x and y and substituting in the equation for z, we have for the adjusted values

$$x = 36° 25' 44.23'',$$
$$y = 90° 36' 22.46'',$$
$$z = 52° 57' 53.31''.$$

Substituting these values of x and y in the original equations we find

$$v_1 = -2.77'', \quad v_2 = -5.54'', \quad v_3 = 3.69'',$$

then by (15), putting $n = 3$ and $q = 2$, we have $r = 7''$, nearly. Now in this case we have

$$\Delta = \begin{vmatrix} 7 & 3 \\ 3 & 5 \end{vmatrix} = 26, \quad \Delta_a = 5, \quad \Delta_b = 7,$$

so that by (17)

$$A_1 = \frac{7}{26}, \qquad B_2 = \frac{5}{26},$$

and therefore by (16) $r_x = 3.5''$, and $r_y = 3.0''$, nearly.

4. **General observation equations.** We have just considered the determination of the most probable values of several unknown quantities as a result of observed values of linear functions of those quantities. It is now proposed to show that the method may be made to serve even if the functions observed are not linear.

Suppose the observation equations to have the form

$$\begin{aligned} f_1(x, y, z) &= M_1 \\ f_2(x, y, z) &= M_2 \\ &\cdots\cdots\cdots \\ f_n(x, y, z) &= M_n \end{aligned} \tag{18}$$

where the functions on the left are known functions, but of any form whatever, provided only that they are such as to possess partial derivatives with respect to all of the variables. Let x_o, y_o, z_o be approximate values of x, y and z, found by solving three of the equations, if not otherwise known. Now write

$$x = x_o + h, \quad y = y_o + k, \quad z = z_o + m,$$

then, by Taylor's theorem, neglecting higher powers of h, k and m,

$$f_1(x, y, z) = f(x_o, y_o, z_o) + h\left[\frac{\partial f_1}{\partial x}\right]_o + k\left[\frac{\partial f_1}{\partial y}\right]_o + m\left[\frac{\partial f_1}{\partial x}\right]_o \quad (19)$$

where the partial derivatives are evaluated at (x_o, y_o, z_o). Here h, k and m are corrections to be added to the approximate values x_o, y_o and z_o, and the equation (19) will be satisfactory if these corrections are small. There are n equations of form (19), and the observation equations (18) may be replaced by n of the form

$$h\left[\frac{\partial f_1}{\partial x}\right]_o + k\left[\frac{\partial f_1}{\partial y}\right]_o + m\left[\frac{\partial f_1}{\partial z}\right]_o = M_1 - f_1(x_o, y_o, z_o). \quad (20)$$

These are linear equations for h, k and m, and normal equations may be formed as before.

Here, also, the method may be extended to functions of any number of unknowns. It may be observed that the general observation equations discussed here include the case of linear equations, and that equations of form (20) may often be substituted for linear observation equations to great advantage.

For example, consider the three angles of the triangle in the example above. Approximate values are the measured values of x and y, and writing

$$x = 36^\circ\ 25'\ 47'' + h, \quad y = 90^\circ\ 36'\ 28'' + k,$$

we have, corresponding to (20)

$$h = 0, \text{ weight } 4,$$

$$k = 0, \text{ weight } 2,$$

$$h + k = -12'', \text{ weight } 3.$$

The normal equations are

$$7h + 3k = -36''$$

$$3h + 5k = -36,$$

and these give for the corrections $h = -2.77$ and $k = -5.54$, and the values of x and y are as found before.

EXERCISES LVI

1. A base line is measured five times with a steel tape reading to hundredths of a foot, and also five times with a chain reading to tenths of a foot, with the following results:

By the tape:	By the chain:
741.17 feet	741.2 feet
741.09 feet	741.4 feet
741.22 feet	741.0 feet
741.12 feet	741.3 feet
741.10 feet	741.1 feet

Find the probable errors of single observations in the two cases and find the proper adjusted value of the length of the line.

2. The length of a line is stated by one party as 683.4 ± 0.3 and by a second party as 684.9 ± 0.3. What is to be inferred from the two results?

3. If six measurements yield a mean whose probable error is 0.02, how many additional measurements of the same precision will reduce the probable error of the mean to 0.004?

4. Ten measurements of the density of a body made with equal precision gave the results

9.662	9.664	9.677	9.663	9.645
9.673	9.659	9.662	9.680	9.654

What is the probable density of the body and what is the probable error of this result?

5. The following measurements were made, with the weights indicated. Find the adjusted value of the quantity measured, and the probable error of this mean,

4.512 weight 2,
4.515 weight 1,
4.507 weight 4,
4.503 weight 3,
4.502 weight 2,
4.511 weight 2,
4.497 weight 3.

6. A gun is fired twelve times, with the same aim, at a vertical screen. The vertical distribution of the shots, shown by distances in feet from the base of the screen of the points of impact, is as follows:

7.5	9	4.5	11.5	6	8.5
13	7	9.5	6.5	10	9

(a) How high is the probable point of aim?

(b) Find the probable vertical deviation of a shot from the point of aim.

(c) Find the probability of hitting a horizontal strip 4 feet wide on the screen: (1) aiming accurately at its center line, (2) aiming accurately at its bottom line.

(d) Assuming that the horizontal dispersion is the same as the vertical, find the probability of hitting a spot four feet square on the screen: (1) with the aim at its center, (2) with the aim at one corner, (3) aim at the middle of one edge.

7. The measured values of four angles at a station are a, b, c and d. The measured value of their sum is a+ b + c + d + e. What are the adjusted values of the angles? What are their probable errors?

8. The following levels were taken to determine the elevations of five points, A, B, C, D and E above O.

A above O = 115.52,	D above C = 632.25
B above A = 60.12,	D above E = 211.01
B above O = 177.04,	E above B = 596.12
C above A = 234.12,	E above C = 427.18
C above B = 171.00.	

Find the adjusted elevations, with their probable errors.

9. Gordon's formula for ultimate strength of a column may be written

$$p = \frac{S}{1 + Tr^2}$$

where S and T are constants, r is the ratio of the length of the column to its least diameter and p is the load in pounds per square inch of cross-section area. Find S and T from the data

p = 34650,	35000,	36580,	37030,
r = 42,	33,	24,	19.5.

10 Assuming $x = 5$ and $y = 10$ as approximate values, determine the best values of x and y from the following observational equations

$$\begin{aligned} x^2 + xy &= 75.2, \\ x + y^2 &= 104.5, \\ 2xy - y^2 &= 0.3, \\ 3x + y &= 24.6. \end{aligned}$$

Determine also the probable errors of the adjusted values of x and y.

11. To determine the dimensions of a right circular cylinder independent measurements were made of length and diameter, and its volume was determined by its loss of weight when submerged in water. The results are:

$$h = 3.225, \quad 2r = 1.375, \quad \pi r^2 h = 14.348.$$

Find h and r and their probable errors: (a) giving the observations equal weight, (b) giving weights of 1, 2 and 10, respectively.

12. In a sine intensity magnetometer, let the pole strength of the bar magnet be P, the distance between its poles d, and the distance from its center to the needle pivot a. If θ is the needle deflection and H the horizontal intensity of the earth's magnetism, the equation connecting these quantities is

$$\frac{P}{H}(2a^2d + d^3) = a^5 \sin\theta.$$

Find the most probable values of P and d, from the following data obtained at a station where $H = 0.1884$ (c. g. s.).

a (c.m.)	θ	a (c.m.)	θ
20	24° 17′	40	1° 49′
25	12 46	45	1 35
30	6 48	50	1 27
35	3 26		

Art. 60. Practical Bearing of Precision

In the two preceding articles we have presented briefly those features of probability and the method of least squares that are most vital in dealing with empirical data. As was indicated in Art. 58, the form adopted for the probability integral cannot be established rigorously.

The assumptions we have made, or others of similar significance, are an unavoidable feature of the discussion. Other assumptions, of different significance but no less plausible, might lead to equations of quite different form. But whatever its lack of rigorous deduction, it should be understood that the equation

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2} \tag{1}$$

and the consequent method of least squares rests upon a genuinely scientific basis. In almost any case where a long series of observations or measurements are made the curve showing the distribution of deviations from the mean is found to be in substantial agreement with the graph of (1) for some value of the parameter h. Being thus in accord with experience, the equation is justified on a purely empirical basis, wholly independent of any theoretical deductions.

The brief discussions of the topics which follow are intended to facilitate correct and convenient application of the method.

1. **Kinds of errors.** There are many kinds of errors, some avoidable, some that may be discovered and corrected, and some that must unavoidably remain to affect the results sought. No classification can be complete, but we may consider the following.

(a) **Persistent errors.** These are errors that have like effect on all the observations in a given series. Such an error will obviously have the same influence on the mean of the set as on each separate observation, but can have no effect on the residuals. Clearly, then, a persistent error renders the mean a less accurate estimate of the true value sought, but this inaccuracy is not evident in the probable error, or measure of precision, which is determined by the residuals. Among persistent errors is the **personal equation** of the observer. Most individuals show a persistent tendency to assign a value too large or too small to observations of any particular type. The time of reaction to a signal is different for different individuals. Such personal tendencies may be carefully evaluated by properly planned tests and their effects may be corrected accordingly. **Instrumental** errors, due to imperfect adjusment or design, are also persistent. They should be evaluated by carefully planned experiments and checks, comparison with other instruments, etc. Corrections are usually made for such errors. Instances of instrumental errors are the "rate" of a clock, the calibration error of a thermometer, the zero circle of a planimeter. Other persistent errors may be inherent in the method employed and are known by the theory of the experiment. Such **theoretical** errors should

be removed by corrections. Correction for variations from standard temperature, barometric pressure, for effects of wind, loss of heat, evaporation, etc., may be of this class.

(b) **Accidental errors.** These are errors that arise in each separate observation independently. It is not possible to foresee them, or to evaluate them, and therefore not possible to correct them. It is from errors of this class that residuals arise, and they determine the precision of the individual observation as well as that of the mean. It is clear, then, that the reliability of an estimate of an unknown quantity is quite distinct from its precision. But if all persistent errors are eliminated, then the precision is a measure of the reliability of the result. If a quantity may be determined by two or more series of observations using independent methods, different instruments, etc., and if the differences between the results obtained are not large in comparison with the probable errors of the separate means, there is strong indication that no persistent errors of harmful magnitude are present.

(c) **Mistakes and large errors.** Occasionally in a series of observations one or two observations are found to differ greatly from the others. In such case there is strong evidence that a **mistake** has been made, such as putting down the wrong figure, reading a scale backward, recording the complementary angle, etc. Such a mistake is not an error in the sense used in discussing precision. Mistakes may usually be avoided by care, and, when they occur, may usually be detected and corrected, or the observation affected may be thrown out. Large errors may appear also as a result of external disturbances, such as a jar to an instrument, presence of foreign matter, a gust of wind, etc. In any case, the large residual should be investigated, and if it may be explained by a mistake or external disturbance, the proper correction should be made. If the large residual can not be sufficiently explained, still the observation involved may be assigned less weight than the others or discarded. Judgment must be used as to what is to be regarded as a large residual. This will, of course, depend on the number of observations and on their precision. Suppose a certain observation to be in question, and let the quantity h be determined for the series. Now

$$P = \frac{2h}{\sqrt{\pi}} \int_0^a e^{-h^2x^2}\, dx$$

is the probability that a residual shall be between $-a$ and a, and $1 - P$ is the probability that a residual will be outside this interval.

In the set of n residuals $n(1 - P)$ is the number of residuals that should be greater numerically than a. It is considered good practice to discard an observation if it gives a residual greater than a, where a is so chosen that $n(1 - P) = \frac{1}{2}$. If this number a, so chosen, is written as $a = br$, the values of b are approximately as follows:

n:	5	10	20	40	80	160
b:	2.4	2.9	3.3	3.7	4.1	4.4

This **criterion for rejection** of observations is not to be followed strictly or blindly, but may be held as an aid to sound judgment. A common practice is to regard 4r as a bound for residuals whatever the number of observations, providing the number exceeds three or four.

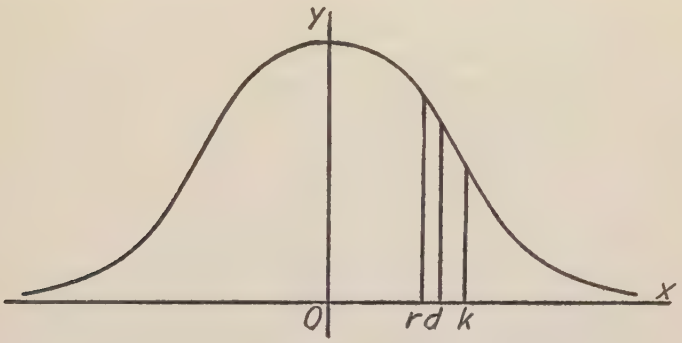


FIG. 111

2. **Precision and significant figures.** Let the graph of equation (1) above be represented in Figure 111. Let the ordinate at r divide the area under the right half of the curve into two equal parts; let the ordinate at d pass through the center of gravity of this area, and let k be the radius of gyration of this area. By setting up and evaluating the necessary integrals it may be shown that

$$d = \frac{1}{h\sqrt{\pi}}, \qquad k = \frac{1}{h\sqrt{2}}, \tag{2}$$

while r is, of course, the probable error defined in Art. 59. Numerical relations are

$$r = 0.8453\ d = 0.6745\ k. \tag{3}$$

The quantity d is called the **average deviation** of the observations from the mean, and k is called the **mean error.** Corresponding to the proba-

ble error r_o of the mean of n observations we have the **average deviation of the mean** and the **mean error of the mean** given respectively by the equations

$$d_o = \frac{d}{\sqrt{n}}, \quad k_o = \frac{k}{\sqrt{n}}. \tag{4}$$

The values of d and k may be computed directly from the residuals for a given set of observations. If the residuals are v_1, v_2, v_3, etc., the obvious equations are

$$d = \frac{|v_1| + |v_2| + \ldots + |v_n|}{n} = \frac{\Sigma|v|}{n}, \tag{5}$$

$$k = \sqrt{\frac{v_1^2 + v_2^2 + v_n^2}{n}} = \sqrt{\frac{\Sigma v^2}{n}}. \tag{6}$$

These expressions are satisfactory for large values of n, but for small values of n they imply a false degree of precision, since they assume the mean to be at the center of the standard curve given by (1). For $n = 1$ these equations must obviously be rejected. It may be shown that equations suitable for all values of n are

$$d = \frac{\Sigma |v|}{\sqrt{n(n-1)}}, \tag{5'}$$

$$k = \sqrt{\frac{\Sigma v^2}{n-1}}. \tag{6'}$$

On account of the greater ease of computing d, and d_o, they are often used in place of r and r_o in indicating precision. Occasionally k and k_o are used, but they offer no important advantages. By equations (2) and (4) it may be observed that k_o is identical with the quantity q that entered the discussion in Art. 59.

Since d and k computed by (5) and (6) or (5′) and (6′) fulfill (3) approximately, it is evident that an approximate relation is

$$0.8453\, \Sigma |v| = 0.6745 \sqrt{n \Sigma v^2}.$$

Making this substitution in the expressions for r found in Art. 59 we obtain the approximate formula

$$r = 0.8453 \frac{\Sigma |v|}{\sqrt{n(n-1)}}, \tag{7}$$

and corresponding approximations are available for r_o and h. These approximations possess the advantage that the sum of the absolute

values of residuals is more easily computed than the square root of the sum of squares of residuals. The coefficient of $\Sigma|v|$ in (7) may be found tabulated as a function of n in physical and mathematical tables.

The value of r_o, d_o or k_o, which is to be recorded as an indication of the precision of a determination, should be written to two significant figures. Occasionally only one significant figure is used if the digit involved is 8 or 9. The determination itself is written to as many significant figures as necessary to include the decimal place occupied by the last digit used to indicate precision. Thus the recorded value 2.718 ± 0.024 indicates that 2.718 is the mean of a series of observations and that r_o, d_o or k_o for the series has been found to have the value 0.024. The values 2983 ± 14, 28700 ± 800, 0.0273 ± 0.0019 are correctly recorded, the first being recorded to four significant figures, the others to three significant figures. If a value is stated as 18300 ± 260 the implication is that the zero following 3 has actually been computed, and that the value lies between 18295 and 18305, while the last zero merely completes the decimal places, the fifth digit being unknown. Examples of improper recording are: 1.27 ± 0.053, 2375.2 ± 16, 305.73 ± 2.14.

It is important to observe proper rules in recording determinations, not only to secure the best economy in preserving the results of precision with least labor, but also to conform to a practice that is generally adopted among scientists and technical scholars.

3. **Propagation of error.** In the present section we consider the precision of a quantity which is computed as a known function of one or more quantities, each quantity being subject to a certain probable error. Based on this consideration are definite rules for computation which conserve precision without unnecessary labor.

(a) **Probable error of a function.** Consider a function of three variables,

$$w = f(x, y, z), \tag{8}$$

and write the total differential

$$dw = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz. \tag{9}$$

Now suppose x, y and z each to be the mean of a set of observations, and let dx, dy and dz represent a certain set of residuals, in corresponding series of observations, then the total differential is the corresponding residual for w. But dx, dy and dz are as likely to be positive as

negative, and to get results we square both sides of (9). Replacing dw, dx, dy and dz by v, v_1, v_2 and v_3, we have

$$v^2 = \left(\frac{\partial f}{\partial x}\right)^2 v_1^2 + \left(\frac{\partial f}{\partial y}\right)^2 v_2^2 + \left(\frac{\partial f}{\partial z}\right)^2 v_3^2 + 2\left[\frac{\partial f}{\partial x}\frac{\partial f}{\partial y} v_1 v_2 + \frac{\partial f}{\partial y}\frac{\partial f}{\partial z} v_2 v_3 + \frac{\partial f}{\partial z}\frac{\partial f}{\partial x} v_3 v_1\right]. \quad (10)$$

Let n_1 denote the number of residuals v_1, n_2 of residuals v_2 and n_3 of residuals v_3, then the number of residuals v, allowing for all combinations of v_1, v_2 and v_3, is $n = n_1 n_2 n_3$, and these n values of v are given by n equations of form (10). Adding these n equations we have

$$\Sigma v^2 = n_2 n_3 \left(\frac{\partial f}{\partial x}\right)^2 \Sigma v_1^2 + n_1 n_3 \left(\frac{\partial f}{\partial y}\right)^2 \Sigma v_2^2 + n_1 n_2 \left(\frac{\partial f}{\partial z}\right)^2 \Sigma v_3^2,$$

the summation of the terms in brackets in (10) being obviously zero.

If we divide this equation by n, we find

$$k^2 = \left(\frac{\partial f}{\partial x}\right)^2 k_1^2 + \left(\frac{\partial f}{\partial y}\right)^2 k_2^2 + \left(\frac{\partial f}{\partial y}\right)^2 k_3^2, \quad (11)$$

where the k's are the mean errors of the respective series of observations according to equation (6). Since the k's are proportional to the d's and the r's, the mean errors in (11) may be replaced by the average deviations or the probable errors.

The discussion given here for a function of three variable is evidently applicable to a function of any number of variables. Certain special cases are worthy of notice. For a **linear function** of several quantities, for example

$$w = ax + by + cz + d$$

the partial derivatives are simply the coefficients and we have

$$r^2 = a^2 r_1^2 + b^2 r_2^2 + c^2 r_3^2. \quad (12)$$

If $a = b = c = 1$ then w is simply a sum, and, if $\bar{r}$ is the largest of the three r's on the right, it is evident that r lies between $\bar{r}$ and $\sqrt{3}\,\bar{r}$.

For the **product** of several factors, say

$$w = xyz,$$

we have

$$r^2 = y^2 z^2 r_1^2 + x^2 z^2 r_2^2 + x^2 y^2 r_3^2,$$

and dividing through by w^2 we obtain the equation

$$\frac{r^2}{w^2} = \frac{r_1^2}{x^2} + \frac{r_2^2}{y^2} + \frac{r_3^2}{z^2}. \tag{13}$$

The terms of this equation are the squares of the probable percentage errors. If $\bar{t}$ is the greatest of the three probable percentage errors on the right then the probable percentage error of w lies between $(\bar{t})$ and $\sqrt{3}\,\bar{t}$. Since the percentage error of the reciprocal of a quantity is the same as that of the quantity itself, equation (13) is equally valid if one or more of the quantities appear as divisors so that w is a **quotient.**

Next consider a **power** of a single quantity, say

$$w = x^m.$$

We now have

$$r^2 = (mx^{m-1})^2 r_1^2$$

and dividing by w^2, and using t and t_1 to denote probable percentage errors, we obtain

$$t^2 = m^2 t_1^2. \tag{14}$$

This equation is valid whether m is positive or negative, integral or fractional.

Finally, consider the **logarithm** of a quantity,

$$w = \log x.$$

Here we have

$$r^2 = \frac{r_1^2}{x^2} = t_1^2, \tag{15}$$

that is, the probable error of the logarithm of a quantity is the probable percentage error of the quantity.

(b) **Computation rules.** The following rules to be followed in computation are evidently justified by equations (12), (13), (14) and (15).

(i) **In adding or subtracting a number of quantities find the probable error (or average deviation) of each, and then retain in each quantity as many places of figures as correspond to the second place of significant figures in the largest probable error (or average deviation).**

In other words, when all numbers are properly arranged in columns for addition or subtraction, the last column of digits to the right that contains a significant figure of every number is retained, while digits to the right of this are superfluous.

(ii) **In multiplication or division determine the probable percentage error (or deviation) of each number, retain enough significant figures in each factor or divisor, and also in the result, to correspond to the greatest probable percentage error (or deviation).**

In other words, all factors and divisors that are subject to error should be expressed to the same number of significant figures, the number of figures to be determined by that factor or divisor having the least number of significant figures; provided that a number whose first digit is small, as 1 or 2, may have one more significant figure than another whose first digit is large, as 8 or 9.

(iii) **In using logarithms for multiplying or dividing, retain as many places in the mantissa of the logarithms as are properly retained in the factors themselves under rule** (ii).

The full advantage of rule (ii) is not secured unless correct methods of multiplication are employed. To illustrate suppose it required to find the circumference of a circle the diameter of which is properly recorded as $d = 27.86 \pm 0.35$. The process is:

Correct method		Common method	
	27.86		27.86
	3.142		3.142
	8358		5572
	2786		11144
	1116		2786
	56		8358
	87.54		8753612
			87.54

In the correct method we multiply by 3 then by 0.1, then 27.9 by 0.04, then 28 by 0.002. Thus the partial products are carried one place beyond that to be retained.

EXERCISES LVII

1. Discuss the two series of observations in problem 1 of Exercises LVI with regard to rejection of observations and the evidence of persistent errors.

2. Give an answer to problem 2 of Exercises LVI.

3. Apply the criterion for rejection of observations to the measurements in problem 4 of Exercises LVI. Also compare the approximate value of r given by (7) with the value previously found by (5) of Art. 59.

4. The altitude and diameter of a certain right circular cylinder are measured with the results:

$$\text{diameter} = 2.542 \pm .014 \text{ in., altitude} = 1.753 \pm .016 \text{ in.}$$

Find the volume of the cylinder, indicating precision in approved form.

5. The maximum speed of a certain motor boat in still water is 28.46 ± 0.24 miles per hour. How fast can the boat go up stream when the current is 4.35 ± 0.06 feet per second? Indicate the precision of the result.

6. In estimating the attendance at a certain convention A estimated the number on the main floor at 2400, B estimated the number in the first gallery at 950, and C estimated the number in the second gallery at 550, thus arriving at an estimated total of 3900. If A, B and C each have a probable percentage error of 15 per cent, what is the probable percentage error of the total?

7. In estimating the acreage of a certain crop a certain state is divided into 9 districts, with an observer in each district. Assuming that the districts plant approximately equal areas and that the observers estimate their respective districts with about equal precision, show that the estimate of the total acreage has a probable percentage error about one-third that of the estimate of each observer for his district.

8. The value of g is to be determined experimentally by use of a pendulum which beats seconds approximately. The formula is

$$g = \frac{\pi^2 L}{t^2}.$$

If the probable percentage error of L is 0.0025 and that of t is 0.0015 what is the probable error in g?

9. A formula from hydraulics connecting head and velocity of flow in long pipes is

$$H = f \frac{L}{d} \frac{v^2}{2g},$$

where f is an empirical constant, L is length of pipe in feet, d is diameter of pipe in feet. If $f = 0.0412$, $L = 7948$ feet, $d = 6$ inches and $v = 10$ feet per second, what is the required value of H? Observe correct rules as to significant figures.

10. The angle of twist of a round shaft under the torque M in terms of its length L and radius r is

$$\theta = \frac{2ML}{\pi r^4 E_s},$$

where E_s is the modulus of elasticity in shear. Given $E_s = 12{,}000{,}000$ (pounds per square inch), $L = 72$ inches, $r = 1\frac{3}{16}$ inches and $M = 3500$ pounds-inches, compute θ, with due regard to economy and precision.

Art. 61. Empirical Equations

An equation expressing a relation among two or more variables is called an **empirical equation** if either the form of the equation or the values of one or more of the constants which enter the equation may be regarded as the result of observation or of experimental data. It is the purpose of the present article to indicate some of the methods most often useful in obtaining such results from given data.

We find it convenient to distinguish between two types of empirical equations. An equation is said to be **rational** in form if by processes of reasoning a form may be derived for the equation, lacking only the values of certain constants. These constants are then said to be **empirical constants.** If no such rational form is known, or if such form is inconvenient, it may be possible and desirable to select an equation of convenient form which expresses, with sufficient accuracy, the relation indicated by the available data. Such an equation is called an **empirical formula.** It should not be understood that the two types of equations are sharply separated. Certain equations, such as Gordon's formula for the strength of long columns used in the exercises above, may be called semi-rational, the form being chosen to satisfy some of the demands of logical development. A basis for such development, however, is Hooke's law, which is itself an empirical formula, and which contains as a constant the experimentally determined modulus of elasticity. The tendency to lose sight of the empirical character of such fundamental and well established relations as Hooke's law serves to render less sharp the discrimination between rational and empirical forms.

If the form of an equation is truly rational it may be used with comparative accuracy over a range considerably wider than that covered by the data used to determine the empirical constants of the equation. But an empirical formula is reliable only for those ranges of the variables for which its accuracy has been tested. Empirical equations are used largely for their convenience in application of mathematical processes to the functions represented. Care should be exercised, however, as to the accuracy of such processes, especially the process of differentiation.

1. **Determination of empirical constants.** The determination of the values to be assigned to the constants which enter an equation is inde-

pendent of the process by which the form of the equation is selected, whether this form be rational or entirely empirical. Let us consider, then, the manner of using empirical data to determine constants, assuming that the form of the equation is known.

Let us suppose that two variables, x and y, are connected by an equation of known form, and that the equation involves several constants, say n. Then by transposing all to the left, the relation may be written

$$f(x, y, c_1, c_2 \ldots c_n) = 0.$$

The empirical data consist of several pairs of corresponding values of x and y, measured, counted, observed, as the case may be. The greater the number of observations the more reliable may be the conclusion drawn. Even a first approximation of the c's requires that n pairs of values of x and y be known, to yield n equations for the n unknown constants. Let n_1 be the number of observed pairs of values of x and y, and suppose n_1 is greater than n. We may plot as rectangular coordinates, each observed pair of values giving a point, (x, y). It should now be possible to draw a smooth curve such that the distance of each point from the curve is within a reasonable limit of error for the observations. Of course the curve should be drawn with attention to the character of curves represented by the equation for assumed values of the constants. Also, the curve should be so drawn that the plotted points are about equally distributed on both sides of the curve. We may now choose n points on this curve, at convenient intervals and well distributed along the curve within the range of the data, and placing the n pairs of coordinates for x and y in the given equation we form n equations which may be solved for the n constants.

In this manner the curve drawn is substituted for the observed data. The value of the method will depend on the care and skill with which the curve is drawn. But since each observation is subject to error, a curve drawn with proper attention to all the observed data should yield n points which determine more probable values of the constants than may be determined by any particular n pairs of observed values.

The accuracy and the ease with which the constants may be determined will, of course, depend on the form of the equation and on the number of unknown constants. If the equation is linear in the constants their determination is direct and certain, while if the equation is of higher degree in the constants, or involves irrational or transcendental

functions of them, solution of the n equations may offer serious difficulty, and may be possible only by approximate methods.

For many scientific data, and for most engineering data, the method we have indicated will prove sufficiently accurate. But if a more refined method for finding the most probable values of the constants is desirable we may resort to the method of least squares. The equation

$$f(x, y, c_1, c_2 \ldots c_n) = 0$$

yields as many different equations as we have observed pairs of values of x and y, and these equations are, in fact, observation equations for the unknown constants. We may form a set of n normal equations and thus find the most probable values of the constants. The great labor involved renders this treatment undesirable unless very precise determinations are required, and unless the data are sufficiently precise to justify the labor.

Certain particular cases are worthy of consideration, both because they are of frequent occurrence and because they yield readily to simple methods.

(a) **Two constants, involved linearly.** In this case our equation may be written

$$C_1 f_1(x, y) + C_2 f_2(x, y) + f_3(x, y) = 0, \qquad (1)$$

or, after dividing by one of the first two functions and transposing,

$$Y = mX + b$$

where Y and X are each a ratio of two functions and m and b replace C_1 and C_2. Now by computing values of X and Y from the given data, we may obtain a series of points (X, Y) which must lie approximately on a straight line. This line may be drawn, with proper attention to all the plotted points, and the slope and y-intercept of this line are the values of m and b. The slope and intercept may be measured directly, or two points of the line may be used to determine the equation of the line in terms of X and Y. It should be clear that the scales for X and Y need not be the same, but in measuring for m and b proper account must be taken of these scales.

(b) **Power curves, logarithmic plotting.** If the equation may be put in the form

$$y = ax^n \qquad (2)$$

where y and x are any functions of the variables observed, and a and n are the constants to be determined, then by taking logarithms of both sides of the equation we have

$$\log y = n \log x + \log a.$$

This now falls under case (a), putting $Y = \log y$, $X = \log x$, $m = n$ and $b = \log a$, and the constants are easily determined. The work here is greatly facilitated by use of logarithmic paper, that is, paper ruled with a variable scale in such manner that the logarithms of x and y may be plotted to scale without reference to a table of logarithms. When the points (x, y) are plotted on this paper the value of a is read in the scale for y at $x = 1$, while the slope n of the line drawn through the points must be actually measured by ordinary means, using a uniform scale.

(c) **Exponential curves, semi-logarithmic plotting.** If the equation has the form

$$y = ae^{mx}, \tag{3}$$

then by taking logarithms we have

$$\log y = m \log e \cdot x + \log a,$$

and again the work falls under case (a), putting $Y = \log y$ and $X = x$, while as before log a is the intercept of the line, and m log e is its slope.

In this case the work is facilitated by use of semi-logarithmic paper, ruled as logarithmic paper, with the variable scale marked on the bottom and left margins, but with the uniform scale marked on the top and right margins. We may plot then, using the natural scale for x and the logarithmic scale for y. Here a is read in the y scale at $x = 0$, and m log e is the measured slope of the line. Since semi-logarithmic paper is made with 1 and 10 of the logarithmic scale corresponding to 0 and 1 of the natural scale, we must consider our logarithms to the base 10.

2. **Determination of form.** When it is desired to derive an equation which may approximately represent a relation between two variables, and when the state of knowledge of this relation affords no basis of choice of form for the equation, the selection of a suitable form must necessarily be an uncertain process. No general method of attack is available, and it is not possible to offer a criterion by which to decide, in advance of tentative trials, which of several modes of procedure is to be most successful. Moreover a given function may be repre-

sented in two or more different forms with nearly equal accuracy, and the choice of form may be determined as much by suitability and convenience for the use intended as by facility and accuracy of derivation. It should be stated, also, that many relations between variables exist for which data are available in tabular form, but for which no empirical formula has been found that is sufficiently accurate and convenient to be useful.

Given the data upon which an empirical formula is to be based, the first step is usually to plot as rectangular coordinates and draw a smooth curve through or near each of the plotted points. The form of this curve may suggest one or more forms for the equation, the coefficients or constants in which may be found as in Section 1 above. No such form should be accepted until, after all constants are determined, the equation is found to agree reasonably closely with all the data.

Since it is usually difficult to recognize curves by the appearance of limited portions of them, while straight lines may always be so recognized, it is very helpful if we may find a transformation of our variables which may transform the curve to a straight line. This will surely be possible if the equation may be represented in one of the particular forms, (a), (b) or (c) in section 1 above. For this purpose we may tabulate in parallel columns the observed values of our variables, say x and y, and numerous functions of these values, such as x^2, y^2, xy, $\frac{x}{y}$, log x, log y, etc. If, now, we are able to select any two columns which, when plotted as rectangular coordinates, yield a set of points that lie approximately on a line, then the form (a) of section 1 is available, and the two constants may at once be found. If we plot the data for x and y on logarithmic paper, this is equivalent to plotting together the columns for log x and log y, while semi-logarithmic paper may be used for x and log y and for y and log x. The constants are found from the logarithmic or semi-logarithmic paper just as in section 1. The number of trials, i.e., the number of different columns that may be tabulated and plotted by pairs, is limited only by the time available and by the patience of the investigator. Care is necessary, however, to avoid repetition. For example, the equation $xy = mx + b$ is equivalent to $y = b\left(\frac{1}{x}\right) + m$, and if one of these forms proves unsatisfactory it becomes folly to try the other.

Semi-logarithmic paper sometimes may be made to serve for the form

$$y - b = ae^{mx} \tag{4}$$

but the value of b must be found by successive trials. If this is a suitable form and if b has its correct value, then log (y − b) and x, plotted as rectangular coordinates by the aid of the special paper, should yield a straight line. If b is too large the graph will be a curve, having the true line as an asymptote. If b is too small the graph will be a curve having the same asymptote, but on the other side of it. A few trials, therefore, should yield a suitable value for b.

In a similar way logarithmic paper may serve for the form

$$y - b = mx^n. \tag{5}$$

Here, of course, (0, b) is a point on the curve. If b has its correct value then log x and log (y − b), plotted as rectangular coordinates, by the aid of the logarithmic paper, should give a straight line. If the value of b is not well chosen the plotted points will lie on a curve, of which the true line, if such exists, is an asymptote. The failure, after a reasonable effort at adjustment, to discover such a line, and thereby to fix b and determine m and n, must be accepted as evidence that the proposed form is not suitable to represent the function.

A form suitable to a wide class of functions is the polynomial

$$y = c_0 + c_1x + \ldots + c_nx^n. \tag{6}$$

Usually n need not be greater than 3 or 4, and the c's may be determined by the use of n + 1 pairs of values of x and y. Of course n is here a positive integer, while in the previous form it may be negative or fractional. This last form is convenient for the application of the method of least squares, being linear in the c's, but if n is large the labor of forming and solving the normal equations is very great.

Inspection of the tabulated data may result in a suggestion of form. If the values for x are in arithmetical progression, and those for y are also, then a straight line form is obvious. If the values for x are in arithmetical progression and those for y are in geometrical progression, then semi-logarithmic plotting must yield a straight line. If the values of x are in arithmetical progression, and if the values of y are denoted by $y_1, y_2 \ldots y_n$, and we form the first differences,

$$y_1' = y_1 - y_2, y_2' = y_2 - y_3, \ldots y'_{n-1} = y_{n-1} - y_n,$$

and the second differences,

$$y_1'' = y_1' - y_2', y_2'' = y_2' - y_3', \ldots, y''_{n-2} = y'_{n-2} - y'_{n-1},$$

and so on, we may arrive at a set of differences that are all approximately equal. In this case the function is of the form (6) and the degree of the polynomial is equal to the order of the differences that are all equal. For example, if the third set of differences are all nearly equal, then the polynomial should be of the third degree.

EXERCISES LVIII

1. The density s of aqueous solutions containing p per cent of calcium chloride at 15° Centigrade was measured as follows:

p.....	2	4	6	8	10	20
s.....	1.02	1.03	1.06	1.07	1.09	1.18
p.....	25	30	35	40		
s.....	1.23	1.29	1.34	1.40		

Determine s as a linear function of p.

2. If the latent heat, L, of steam at a temperature of t° Centigrade is found as follows, determine a linear relation between L and t.

t.....	75	90	100	115	125
L.....	554	544	536	526	519

3. Determine a and b so that the equation $y = a + bx^2$ shall fairly represent the law indicated by the data:

x.....	19	25	31	38	44
y.....	1900	3230	4900	7330	9780

4. The following quantities are supposed to follow a law of the form $y = ax^n$. Test this, and find the best values of a and n.

x.....	4	7	11	15	21
y.....	28.6	79.4	182	318	589

5. The following table gives the loss of power E due to magnetic hysteresis, for different values of the magnetic induction B in a certain transformer core.

B.....	2000	4000	6000	8000	10000
E.....	2869	8700	16660	26370	37660

Determine a formula of the form $E = aB^n$.

6. Test the following data for a law of the form $y = ae^{mx}$, and find the values of the constants.

x.....	2.701	2.870	3.258	3.681	3.892
y.....	3.86	4.20	5.10	6.30	7.00

7. The following data give θ, the difference of temperature of a cooling body and the surrounding air, at time t seconds.

t.....	0	3.45	10.85	19.30	28.80	40.10	53.75	70.95
θ.....	19.9	18.9	16.9	14.9	12.9	10.9	8.9	6.9

Newton's law of cooling is $\theta = \theta_1 e^{-mt}$. Test the law and find θ_1 and m.

8. The current A (amperes) in an electric arc at potential difference V (volts) was measured as follows:

A.....	1.96	2.46	2.97	3.45	3.96	4.97	5.97
V.....	50.25	48.7	47.9	47.5	46.8	45.7	45.0

A.....	6.97	7.97	9.00
V.....	44.0	43.6	43.5

Determine a formula involving two empirical constants. (Compute products VA as a new variable.)

9. The temperature θ at depth x in a certain artesian well was measured as follows:

x.....	28	66	173	248	296	400	505	548
θ.....	11.71	12.90	16.40	20.00	22.20	23.75	26.45	27.70

Find a law of the form $\theta = a + bx + cx^2$.

10. The electrical conductivity C of glass at temperature θ is

θ.....	58	86	148	166	188	202	210
C.....	0	.004	.018	.029	.051	.073	.090

Find an empirical formula connecting θ and C.

11. The volume v (cubic feet) of a pound of saturated steam at pressure p (pounds per square inch) is given as follows:

p.....	10	20	30	40	50	60
v.....	37.80	19.72	13.48	10.29	8.34	6.62

Find an empirical formula.

12. Find a formula of the form $y = b + ae^{mx}$ for the following data:

x..	4	5	6	7	8	9	10	11	12	13
y..	20.1	27.8	35.0	41.5	47.6	53.0	58.1	62.7	66.8	70.0

13. The theoretical law for electrical current i at time t has the form $i = I(1 - e^{-kt})$. Test the form and find I and k from the data

t.....	o	0.2	0.4	0.6	0.8	1.0
i.....	o	0.075	0.20	0.34	0.45	0.49

14. The following data pertain to a lens; p is the distance of the object, p′ the distance of its image.

p.....	320	240	180	140	120	100	80	60
p′.....	21.35	21.80	22.50	23.20	23.80	24.60	26.20	29.00

Find a law connecting p and p′.

BOOKS FOR SUPPLEMENTARY READING

Practical Least Squares, O. M. Leland (McGraw-Hill, 1921).
The Calculus of Observations, Whitaker and Robinson (Blackie and Son, 1924).
Combinations of Observations, D. Brunt (Cambridge University Press, 1917).
Theory of Errors and Least Squares, L. D. Weld (Macmillan, 1916).
Method of Least Squares, Mansfield Merriman (Wiley, 1913).
Method of Least Squares, D. P. Bartlett (Boston, 1900).
Theory of Errors and Method of Least Squares, W. W. Johnson (Wiley, 1912).
An Introduction to the Theory of Statistics, S. U. Yule (Charles Griffin and Company, 1919).
Discussion of the Precision of Measurements, S. W. Holman (Wiley, 1904).
Engineering Mathematics, C. P. Steinmetz (McGraw-Hill, 1911), Chapters VI and VII.
A Course in Practical Mathematics, F. M. Saxelby (Longmans, 1910), Chapter VIII.
Graphical and Mechanical Computation, Joseph Lipka (Wiley, 1918), Chapters VI and VII.
Interpolation, J. F. Steffensen (Williams & Wilkins, 1927).

CHAPTER XIV

ORDINARY DIFFERENTIAL EQUATIONS, FIRST ORDER

Art. 62. Meaning of Differential Equations

In the study of relations that may exist among variables it is often desirable to make use of other relations that may be expressed among these variables and the derivatives of some of them with respect to others. An equation expressing a relation involving one or more derivatives is called a **differential equation.** We have been making use of certain simple differential equations. When a variable is expressed by an integral, as

$$y = \int f(x)\, dx,$$

the equation is equivalent to the differential equation

$$\frac{dy}{dx} = f(x),$$

and leads to a relation expressing y as a function of x and involving a constant of integration. In such simple cases the desired relation is found at once by integration. The theory of differential equations has for its purpose the treatment of equations which involve derivatives and such that it is not possible to integrate until certain transformations have been applied. The relation among the variables, free from derivatives, which results from the transformations and integrations applied to the differential equation, is called the **solution** of the differential equation.

1. **Classification.** Equations involving only one independent variable, so that the derivatives present are of the ordinary kind, are called **ordinary differential equations,** while equations which involve two or more independent variables and partial derivatives with respect to one or more of them are called **partial differential equations.** The present chapter and the two following chapters are concerned wholly with the former class of equations.

The **order** of a differential equation is that of the derivative of highest order present in the equation. For example, an equation involving

derivatives of the third order and of lower order, but none of order higher than the third, is a differential equation of the third order.

The **degree** of a differential equation is that of the highest power of the derivative of highest order present. The equation

$$x^3 \frac{d^2y}{dx^2} + x\left(\frac{dy}{dx}\right)^2 + 3y^4 = 0$$

is of second order and of the first degree, while

$$x\left(\frac{dy}{dx}\right)^2 + 3y - 4x^3 = 0$$

is of the first order and second degree. Of course the degree of an equation pertains to the form it takes when rational and integral with respect to all derivatives which occur.

No general method of solving differential equations is possible. Like integration, the solving of differential equations is an indirect process, and only certain types of equations yield to treatment. We find space here for only the more common and more important types, the appropriate theory and method being indicated for each type. To characterize these types of equations we shall use the properties of order and degree, and other properties to be defined as occasion arises.

2. **Equation and solution.** In order to better understand the indirect process of solving a differential equation, we shall first consider the direct process of deriving the differential equation when its solution is known. As a simple example of an indirect process we may take the integral

$$y = \int 3x^2\, dx,$$

from which we obtain

$$y = x^3 + c. \tag{1}$$

Applying to this the direct process of differentiation we have

$$\frac{dy}{dx} = 3x^2. \tag{2}$$

Equation (2) is a differential equation, and (1) is its solution.

Again, consider the equation

$$y = ae^{2x} + b \sin x, \tag{3}$$

in which a and b are parameters, so that the equation represents a two-parameter family of curves. Taking first and second derivatives we have

$$\frac{dy}{dx} = 2ae^{2x} + b\cos x, \tag{4}$$

$$\frac{d^2y}{dx^2} = 4ae^{2x} - b\sin x. \tag{5}$$

Eliminating a and b from equations (3), (4) and (5), we have

$$(\cos x - 2\sin x)\frac{d^2y}{dx^2} + 5\sin x\frac{dy}{dx} - (4\cos x + 2\sin x)\,y = 0. \tag{6}$$

This second order differential equation has (3) as a solution.

In general, given an equation

$$f(x, y, a_1, a_2, \ldots . a_n) = 0, \tag{7}$$

involving n arbitrary constants, we may form n additional equations by differentiation and eliminate the n constants from the n + 1 equations, and thus obtain a differential equation of the nth order

$$\phi\left(x, y, \frac{dy}{dx}, \ldots . \frac{d^ny}{dx^n}\right) = 0. \tag{8}$$

Equation (7) is known as the **general solution** or **complete primitive** of the differential equation (8). We see that the general solution has a number of arbitrary constants corresponding to the order of the differential equation. If some or all of these constants are given particular values the equation is still a solution, but is no longer the general solution, and is called a **particular** solution.

In the example (3) above, let us choose a particular solution so that when x = 0 we have y = 2. This requires a = 2, and the particular solution is

$$y = 2e^{2x} + b\sin x. \tag{3'}$$

If we further require that the slope at x = 0 shall be zero, we must have b = −4, and the solution becomes

$$y = 2e^{2x} - 4\sin x. \tag{3''}$$

It would be impossible to obtain the differential equation (6) from (3′) or (3″). The equation (3′) is the primitive of a differential equation

of the first order, and (3″) is not the primitive of any differential equation, since it contains no arbitrary constant.

By this illustration we recognize that the general solution or complete primitive is the equivalent of the differential equation with respect to the information it may yield or imply, while a particular solution pertains to the differential equation together with certain additional facts. These additional facts are known as **initial conditions.**

The **test of a solution** is to take the expressions for y and the necessary derivatives as derived from the solution and substitute them in the differential equation, and thus verify the truth of the equation. For example, from (3″) we have

$$y = 2e^{2x} - 4 \sin x,$$

$$\frac{dy}{dx} = 4e^{2x} - 4 \cos x,$$

$$\frac{d^2y}{dx^2} = 8e^{2x} + 4 \sin x,$$

and substituting these expressions in (6) we obtain an identical relation in x.

EXERCISES LIX

1. Show that the equation of any straight line whatever is a solution of the differential equation $\frac{d^2y}{dx^2} = 0$.

2. Find the differential equations whose complete primitives are:

(a) $y^2 = cx$; (d) $y = a \cos 2x$;

(b) $y^2 = c_1x + c_2$; (e) $y = ae^{bx}$;

(c) $y^2 = cx + c^2$; (f) $y = ae^{2x} + be^{-2x}$.

3. Show that the equation of the family of circles of radius 4,

$$(x - h)^2 + (y - k)^2 = 16,$$

is the complete primitive of the differential equation obtained by setting the formula for radius of curvature equal to 4.

4. Given the differential equation

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 12y = 24x,$$

determine which of the following are solutions.

(a) $y = -\frac{1}{6} - 2x$; (c) $y = 8e^{-4x} + \frac{1}{6} - 2x$;

(b) $y = 4e^{3x} - 2e^{2x}$; (d) $y = ae^{3x} + be^{-4x} - 2x - \frac{1}{6}$.

5. Find for what values of a, b and c the equation

$$y = a + bx + cx^2$$

is a solution of the differential equation in problem 4.

6. Show by the differential equation of the family of parabolas

$$y^2 = 8(x + h)$$

that the subnormal of such parabolas has the same value at all points.

7. Given the parabola $y^2 = 16x$ and the family of lines tangent to it, $y = mx + \frac{4}{m}$, write the differential equation of the family of lines and show that the equation of the parabola is a solution of it, as well as any tangent line.

Art. 63. Separation of Variables

In this article and the next we shall consider only differential equations of the first order and first degree. Since the first derivative is the ratio of differentials, any such equation may be written in the form

$$M\,dx + N\,dy = 0, \tag{1}$$

where M and N are functions of x and y. In general we are not able to integrate directly, since y, an unknown function of x, is contained in M, and x, an unknown function of y, is contained in N. In other words, we cannot integrate because the integrand is not expressed as a function of the variable of integration. But if M is a function of x only and N a function of y only the variables are said to be **separated** and we may integrate at once. For example, given

$$x^3dx - \cos y\,dy = 0,$$

we have immediately the solution

$$\frac{x^4}{4} - \sin y = C.$$

in which C is the arbitrary constant.

1. **Variables directly separable.** If M and N may each be separated into two factors, one factor of each being free from x and the other free from y, then by dividing the equation by the factor of M which contains y and the factor of N which contains x we may separate the variables. As an illustration, the equation

$$(x - xy)\,dx + ye^x\,dy = 0$$

when divided by $(1 - y)\,e^x$ becomes

$$xe^{-x}\,dx - \frac{y\,dy}{y - 1} = 0,$$

which may be integrated at once.

2. **Homogeneous equations.** Equations of the type of (1) above are called homogeneous if M and N are homogeneous functions of x and y of the same degree. Now if a function is homogeneous of degree n in x and y, then if we put $y = vx$ it contains x^n as a factor and the other factor is a function of v only. Since y is an unknown function of x, in making such a substitution we must regard v as an unknown function of x, and writing $dy = vdx + xdv$ we have a new differential equation in x and v, and the variables are separable as in 1. An example will make the method clear. Take

$$(y^2 - xy)dx + x^2dy = 0.$$

Putting $y = vx$ and $dy = vdx + xdv$ we have, after dividing out x^2,

$$(v^2 - v)dx + vdx + xdv = 0,$$

or, separating the variables,

$$\frac{dx}{x} + \frac{dv}{v^2} = 0,$$

whence the solution

$$\log x - \frac{1}{v} = C, \text{ or } x = C_1 e^{\frac{1}{v}},$$

where $C_1 = e^C$. Now putting $v = \frac{y}{x}$, the solution becomes

$$x = C_1 e^{\frac{x}{y}}.$$

3. **When M and N are linear.** If M and N are of the first degree in x and y, then the equations $M = 0$ and $N = 0$ represent straight lines. If we shift the origin to the point of intersection of these lines, the new expressions for M and N will have no constant terms and will therefore be homogeneous, so that the method of 2 applies. This type is illustrated by the equation

$$(4x + 3y + 1)dx + (x + y + 1)dy = 0.$$

The lines $4x + 3y + 1 = 0$, $x + y + 1 = 0$ meet at $(2, -3)$ and writing $x = x' + 2$, $y = y' - 3$ the equation becomes

$$(4x' + 3y')dx' + (x' + y')dy' = 0.$$

Now putting $y' = vx'$ and dividing out x',

$$(4 + 3v)\,dx' + (1 + v)\,(vdx' + x'dv) = 0,$$

or

$$\frac{dx'}{x'} + \frac{(1 + v)dv}{(2 + v)^2} = 0,$$

whence the solution

$$\log x' + \frac{1}{2 + v} + \log (2 + v) = C,$$

or, writing $x' = x - 2$, $v = \frac{y'}{x'} = \frac{y + 3}{x - 2}$, we have finally

$$\log (2x + y - 1) = C - \frac{x - 2}{2x + y - 1}.$$

4. **Simple applications.** While we have considered only a few simple types of differential equations we are in position to consider some of the simpler applications which may involve these types. It should be observed, however, that the kind of problem here presented may just as well lead to many other types of differential equations, some of which we shall consider in later articles.

Certain geometrical properties of a curve may be stated in terms of a quantity which involves the first derivative. Instances of such quantities are the subtangent, $y\frac{dx}{dy}$ the subnormal, $y\frac{dy}{dx}$, the intercepts of the tangent, $x - y\frac{dx}{dy}$ and $y - x\frac{dy}{dx}$, the intercepts of the

normal, $x + y\frac{dy}{dx}$, and $y + x\frac{dx}{dy}$. These equations all involve coordinates of the point on the curve and the slope at that point. In polar coordinates we have the tangent of the angle between the curve and the radius vector, $\rho\frac{d\theta}{d\rho}$, the polar subtangent $\rho^2\frac{d\theta}{d\rho}$, the polar subnormal, $\frac{d\rho}{d\theta}$. If we are asked for the curve whose tangent always intercepts the x-axis at (2x, 0) where x is the abscissa of the point of tangency, we have

$$x - y\frac{dx}{dy} = 2x, \text{ or } \frac{dx}{x} + \frac{dy}{y} = 0,$$

from which $xy = C$. This is a family of hyperbolas, any one of which is a curve which has the required property. If we ask for a curve that cuts every radius vector at a given angle, say ψ, then we have

$$\rho\frac{d\theta}{d\rho} = \tan\psi, \text{ or } \cot\psi\, d\theta = \frac{d\rho}{\rho},$$

so that $\theta\cot\psi = \log\rho + C$, and this may be written

$$\rho = C_1 e^{\theta\cot\psi}.$$

This equation represents a family of equiangular spirals, any one of which cuts the radius vector at the given angle.

Equation (1) may be written

$$\frac{dy}{dx} = -\frac{M}{N}, \tag{1}$$

and, since M and N are functions of x and y, for any given point (x, y) the slope is determined by this equation. The solution of this differential equation is a family of curves, called **integral curves,** and each curve has, at every point (x, y) on it, the slope given by this equation (1). If we assign to the slope the negative reciprocal of the value in (1) we obtain a differential equation,

$$\frac{dy}{dx} = \frac{N}{M}, \tag{2}$$

the solution of which is a family of curves meeting everywhere at right angles the curves from (1). The integral curves of (2) are called the **orthogonal trajectories** of the integral curves of (1). Let us find the

orthogonal trajectories of the parabolas $y^2 = 4px$, and, in particular, let us find that one which passes through (2, 4). Taking the derivative

$$2y\frac{dy}{dx} = 4p$$

and eliminating p we have

$$\frac{dy}{dx} = \frac{y}{2x}.$$

The differential equation of the required trajectories is

$$\frac{dy}{dx} = -\frac{2x}{y},$$

from which $y^2 + 2x^2 = C$. This is a family of ellipses, and the particular one going through the point (2, 4) is

$$y^2 + 2x^2 = 24.$$

EXERCISES LX

1. Solve the differential equations:

(a) $(y - 1)dx - (x + 1)dy = 0$, (c) $\sqrt{1 - y^2}\,dx + \sqrt{1 - x^2}\,dy = 0$,

(b) $\sin\theta\ dr + r\cos\theta\ d\theta = 0$, (d) $dy + y\sin x\,dx = \sin x\,dx$.

2. Solve the following:

(a) $(y - x)dy + ydx = 0$, (c) $(x^2y + y^3)dx - x^3dy = 0$,

(b) $(x^2 + y^2)dx = 2xy\,dy$, (d) $(3x^2 - y^2)dy = 2xy\,dx$.

3. Solve the following:

(a) $(2x - y + 1)dx + (x + 2y + 3)dy = 0$,

(b) $(2x + y - 1)dx + (4x + 2y + 4)dy = 0$.

Remark: Here the two lines whose intersection is sought are parallel so the method fails. Put $2x + y = t$, then $dy = dt - 2dx$.

4. Find the curves which have the subtangent constant.

5. Find the curve through (2, 8) which has the subnormal constantly equal to 16.

6. Find the polar equation of the curves for which the polar subnormal is equal to $\rho\sin\theta$.

7. Find the orthogonal trajectories of the family of circles $x^2 + y^2 = r^2$.

8. Find the orthogonal trajectories of the circles $x^2 + y^2 + cx = 0$.

9. Determine the curves such that the angle between the radius vector and the tangent is one-half the vectorial angle.

Art. 64. Exact Differential Equations

A differential equation of the first order and first degree, as

$$M\,dx + N\,dy = 0, \tag{1}$$

has for a general solution, or complete primitive, an equation in x and y and an arbitrary constant C. This primitive may be written in the form

$$f(x, y) = C, \tag{2}$$

and therefore a form of the differential equation is

$$\frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy = 0. \tag{3}$$

The expression here equated to zero is an **exact differential,** and an equation in the form of (3) is called an **exact differential equation.** Now equation (3) determines the ratio of the differentials dy and dx, so that any differential equation, as (1), which has the same solution as (3) must differ from (3) at most by a factor, say $\phi(x, y)$. Given a differential equation (1), therefore, for which there is a solution (2), we see that there is a factor $\phi(x, y)$ such that

$$M\,\phi(x, y) = \frac{\partial f}{\partial x},\ N\,\phi(x, y) = \frac{\partial f}{\partial y}. \tag{4}$$

Such a factor, $\phi(x, y)$ is called an **integrating factor,** and multiplying by it gives an exact differential equation. The importance of this lies in the fact that an exact differential may be integrated without separating the variables. The expressions

$$\int \frac{\partial f}{\partial x}\,dx,\ \int \frac{\partial f}{\partial y}\,dy,$$

where the first regards y as constant and the second regards x as constant, should each yield the function f(x, y), except that the former will not give terms containing y only and the latter will not give terms containing x only. The required function f(x, y) may be obtained, then, by taking one of these integrals plus such additional terms as arise from the other integral.

Take, for example, the primitive

$$x^4 + x^2y^2 - 4y^4 = C.$$

This gives the exact differential equation

$$(4x^3 + 2xy^2)dx + (2x^2y - 16y^3)dy = 0,$$

and if we were seeking the solution of this we should have considered

$$\int (4x^3 + 2xy^2)dx = x^4 + x^2y^2 + C,$$

$$\int (2x^2y - 16y^3)dy = x^2y^2 - 4y^4 + C.$$

Taking the former, together with the term $-4y^4$ which occurs in the latter and not in the former, we have the required solution. It is evident that in the latter integration it is sufficient to consider only those terms that are free from x, since terms containing x must appear in the former integral.

If an equation in the form (1) is an exact differential equation we shall have

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}, \tag{5}$$

for, since M and N are partial derivatives of f(x, y) with respect to x and y respectively, equation (5) is equivalent to the equation

$$\frac{\partial^2 f}{\partial y\, \partial x} = \frac{\partial^2 f}{\partial x\, \partial y},$$

and this is true for any function f(x, y) excepting certain cases of discontinuity, with which we need not be concerned. Equation (5) is known as the **test for exactness** of a differential equation.

It should be noted that any differential equation in which the variables are separated is an exact differential equation, and the separation of the variables by division is, in fact, the use of an integrating factor. The method and point of view of Art. 63, however, seem to gain simplicity by treatment independent of the present article.

1. **Integral factors by inspection.** It is not possible to state any general rule or method for finding integrating factors. Experience, guided by the few suggestions which follow, should lead to a reasonable degree of skill.

(a) The expression x dy − y dx is made an exact differential by dividing by x^2, y^2, xy, $x^2 - y^2$ or $x^2 + y^2$. If an equation can be arranged so as to involve this expression, division by one of these five quantities may be effective. For example given

$$(y^2x + y)dx - x\,dy = 0$$

we may rearrange

$$y^2x\,dx + (y\,dx + x\,dy) = 0$$

and division by y^2 makes the equation exact.

(b) A factor of the form $x^m y^n$ is effective if M consists of two terms of the form $ax^h y^k$, and N consists of two terms formed by multiplying the terms of M by $\frac{bx}{y}$, where b is any constant, and may be different for the two terms. Example:

$$(x^2y^2 - 4xy)dx + (3x^3y - 2x^2)dy = 0.$$

Multiplying by $x^m y^n$ and applying the test for exactness

$$\frac{\partial}{\partial y}(x^{m+2}\,y^{n+2} - 4x^{m+1}\,y^{n+1}) = \frac{\partial}{\partial x}(3x^{m+3}\,y^{n+1} + 2x^{m+2}\,y^n),$$

$$(n+2)x^{m+2}\,y^{n+1} - 4(n+1)x^{m+1}\,y^n = 3(m+3)x^{m+2}\,y^{n+1} + 2(m+2)x^{m+1}\,y^n.$$

This is satisfied if

$$n + 2 = 3(m + 3),\ -4(n + 1) = 2(m + 2)$$

that is, if $m = -\frac{18}{7}$ and $n = -\frac{5}{7}$.

(c) If xM and yN are functions of the product xy, then division by $xM - yN$ will make the equation exact. This fails if $xM = yN$, but in that case the solution is clearly $xy = C$.

(d) If, in applying the test for exactness, we find that $\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} = N\,f(x)$, where $f(x)$ is any function free from y, then multiplication by $e^{\int f(x)dx}$ will make the equation exact.

(e) If $\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} = -M\,f(y)$, where $f(y)$ is any function free from x, then multiplication by $e^{\int f(y)dy}$ will make the equation exact.

2. **Linear equation of first order.** An equation is said to be **linear** if the dependent variable and its derivatives enter linearly. The general form of the linear equation of the first order is

$$\frac{dy}{dx} + Py = Q, \tag{6}$$

where P and Q are any functions of x. While the treatment of this type is a case under the paragraph above, the extreme importance of the type justifies separate consideration. Rewriting we have

$$(Py - Q)dx + dy = 0, \tag{6'}$$

so that $M = Py - Q$ and $N = 1$. Following suggestion (d) above, we have

$$\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} = P - 0 = N\,f(x),$$

where $f(x) = P$. An integrating factor is $e^{\int P dx}$. Multiplying by this,

$$e^{\int P dx}(Py - Q)dx + e^{\int P dx}\,dy = 0.$$

This is readily shown to be exact, and the solution is

$$ye^{\int P dx} = \int Qe^{\int P dx}dx + C. \tag{7}$$

On account of the frequency with which this type occurs it is recommended that equations (6) and (7) be kept in mind, and that given a differential equation of form (6) the integrating factor $e^{\int P dx}$ be at once simplified and applied by use of (7). Consider the example

$$\frac{dy}{dx} + \frac{1-x}{x}y = e^x.$$

Here the integrating factor is

$$e^{\int \frac{1-x}{x}dx} = e^{\log x - x} = xe^{-x},$$

and applying (7) we have

$$xye^{-x} = \int e^x\, xe^{-x}\, dx = \int x\, dx = \frac{1}{2}x^2 + C,$$

and finally the solution is $y = e^x\left(\frac{x}{2} + \frac{C}{x}\right)$.

3. **Extended linear equation.** We treat here a type of equation sometimes called Bernoulli's equation, after James Bernoulli, and which is readily reduced to a linear equation. The type is

$$\frac{dy}{dx} + Py = Qy^n, \tag{8}$$

where n is any number. Divide by y^n, then

$$y^{-n}\frac{dy}{dx} + Py^{1-n} = Q. \tag{8'}$$

Now put $v = y^{1-n}$ and $\frac{dv}{dx} = (1 - n)y^{-n}\frac{dy}{dx}$ then after multiplying by $1 - n$ the equation becomes

$$\frac{dv}{dx} + (1 - n)Pv = (1 - n)Q,$$

which is in the form (6).

EXERCISES LXI

1. Show that the following equations are exact, and write their solutions:

(a) $x^2dx + (y^2 - 1)dy = 0$;
(b) $(x + y)dx + (x - y^2)dy = 0$;
(c) $\cos y\, dx + (y - x \sin y)dy = 0$.

2. Find integrating factors for the following and write the solutions:

(a) $(x^3 - y)dx + xdy = 0$;
(b) $(xy + y^3)dx - (4x^2 + 3xy^2)dy = 0$;
(c) $(3x^2 + y)dx + (x^3y + xy^2 + x)dy = 0$.

3. Solve the following:

(a) $\frac{dy}{dx} - \frac{2y}{x + 1} = (x + 1)^3$;

(b) $\frac{dy}{dx} + y \cos x = \sin 2x$;

(c) $\frac{dy}{dx} + \frac{y}{x} = y^3$;

(d) $\frac{dr}{d\theta} + r\theta = r^2\theta^3$.

4. The speed v with which a certain body moves is subject to the law

$$\frac{dv}{dt} + kv = a \sin t.$$

If $v = 0$ when $t = 0$, express v as a function of t.

5. The electric current in a circuit obeys the law

$$L\frac{di}{dt} + Ri = E,$$

where i is current and t is time, and the other notations indicate constants. Express i as a function of t, putting $i = i_0$ when $t = 0$.

6. In example 5 replace E by E sin t, and solve, noting analogy with 4.

7. If a belt is stretched around a pulley the normal pressure per unit length between the surfaces is T/a, where T is the tension on the belt and a is the radius of the pulley. For an element of length ds the force is $T/a\ ds = T\ d\theta$, where $d\theta$ is the angle subtended by ds (Fig. 112). If f is the coefficient of friction, the corresponding change in tension, when ready to slip, is

$$dT = f\ T\ d\theta.$$

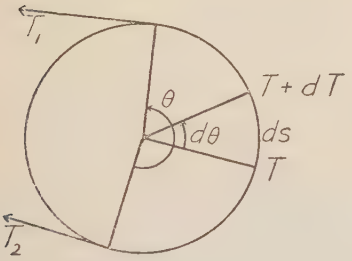


Fig. 112

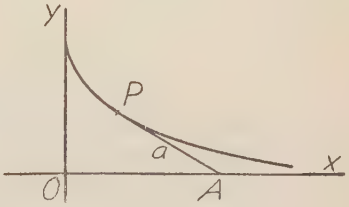


Fig. 113

Show that the ratio of tensions on the belt is limited by the equation

$$\frac{T_1}{T_2} = e^{f\theta}.$$

8. The tractrix, sometimes called the curve of pursuit, is described by a point P drawn by the cord PA of constant length a as the point A moves along OX (Fig. 113). Clearly it is the curve of constant tangent length a. Find its equation. (See exercise 2, II.)

Art. 65. Equations of Higher Degree

In dealing with equations involving higher powers of the derivative it is convenient to use a single letter to denote the derived function. Let $p = \frac{dy}{dx}$, then any equation of the first order, whatever its degree, may be written as

$$f(x, y, p) = 0. \qquad (1)$$

1. **The general solution.** We give here three methods for finding the general solution of an equation of type (1), based respectively on solving equation (1) for p, for y and for x. A choice of methods is important, since, in certain cases, one method may be much more convenient than another.

(a) **Solve for p.** Regarding (1) as an algebraic equation for p, we may solve for p and obtain several values, equal in number to the degree of the equation. Let these values be

$$p = f_1(x, y), \quad p = f_2(x, y) \ . \ . \ . \ . \tag{2}$$

Each of these is a differential equation of the first order and first degree, and may be solved by the methods of previous articles. Let the several solutions be

$$F_1(x, y, c) = 0, \quad F_2(x, y, c) = 0, \ . \ . \ . \ . \ , \tag{3}$$

then the general solution may be written by equating to zero the product, as

$$F_1(x, y, c)\, F_2(x, y, c) \ . \ . \ . \ . = 0. \tag{4}$$

The constants c which occur in these factors must not be regarded as independent, for then equation (4) would be the primitive of a differential equation of higher order than the first, and not of type (1). If these constants c are all the same it may be shown that elimination of c between the derivative of (4) and $F_1(x, y, c) = 0$ gives $p = f_1(x, y)$, and similarly for F_2 and f_2, etc.

As an example take the second degree equation

$$yp^2 + (x - x^2y)p - x^3 = 0.$$

Solving for p we have the two values

$$\frac{dy}{dx} = p = -\frac{x}{y}, \qquad \frac{dy}{dx} = p = x^2,$$

and the solutions are

$$x^2 + y^2 + c = 0, \quad 3y - x^3 + c = 0,$$

or combining into a single equation, the general solution is

$$(x^2 + y^2 + c)\,(3y - x^3 + c) = 0.$$

(b) **Solve for y.** It may be convenient to solve equation (1) for y, expressing y as a function of x and p, say

$$y = \phi(x, p). \tag{5}$$

Now differentiating with respect to x, and remembering that p is a function of x, we have

$$p = \frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial p}\frac{dp}{dx}. \tag{6}$$

This is a differential equation of the first order and first degree in p and x, and may be solved by the methods of previous articles. Let the solution be

$$F(x, p, c) = 0, \tag{7}$$

then elimination of p between (5) and (7) will give the general solution of (1).

Take the equation $p^2 = x - y$. Solving for y we have $y = x - p^2$, and taking the derivative,

$$p = 1 - 2p\frac{dp}{dx},$$

or

$$\frac{pdp}{1 - p} = \frac{dx}{2},$$

and integrating,

$$-p - \log(p - 1) = \frac{x}{2} + C.$$

Now putting $p = \sqrt{x - y}$, we have the solution

$$-\sqrt{x - y} - \log(\sqrt{x - y} - 1) = \frac{x}{2} + C.$$

In some cases the actual elimination of p may prove extremely difficult or awkward. There is no objection to regarding equations (5) and (7), or (1) and (7), together as the solution, c being the arbitrary constant and p serving as a parameter.

A special case is here worthy of attention. If (5) should take the form

$$y = px + \phi(p), \tag{5'}$$

then differentiation gives

$$p = p + x\frac{dp}{dx} + \frac{d\phi}{dp}\frac{dp}{dx}, \tag{6'}$$

and this is satisfied if $\frac{dp}{dx} = 0$, so that (7) takes the form

$$p = c. \tag{7'}$$

The general solution, therefore, is obtained by putting $p = c$ in (5'),

$$y = cx + \phi(c).$$

An equation of the form (5') is known as **Clairaut's equation,** and as soon as the type is recognized the solution may be written without intermediate steps.

(c) **Solve for** x. If it is convenient to solve equation (1) for x we have

$$x = \phi(y, p). \tag{8}$$

Now regarding x and p as functions of y we may take derivatives with respect to y,

$$\frac{dx}{dy} = \frac{1}{p} = \frac{\partial\phi}{\partial y} + \frac{\partial\phi}{\partial p}\frac{dp}{dy}. \tag{9}$$

This is a differential equation of first order and first degree in p and y, and if we write the solution

$$F(y, p, c) = 0 \tag{10}$$

the general solution is given by (8) and (10) or by (1) and (10) either by elimination of p or by retaining the two equations and regarding p as a parameter. Using again the example $p^2 = x - y$ we have $x = y + p^2$, from which

$$\frac{1}{p} = 1 + 2p\frac{dp}{dy},$$

or

$$\frac{p^2 dp}{1 - p} = \frac{dy}{2}.$$

The solution is

$$-\frac{p^2}{2} - p - \log(p-1) = \frac{y}{2} + C,$$

and putting $p = \sqrt{x-y}$ we have the same solution as before.

2. **The singular solution.** Let us consider the differential equation

$$f(x, y, p) = 0 \tag{1}$$

and its general solution

$$F(x, y, c) = 0. \tag{11}$$

Equation (11) may be from form (3) above, or it may come from (1) and (7) or (1) and (10) by elimination of p. This equation (11) represents a family of curves, c being the parameter of the family, and such a family may have an envelope; a curve, every point of which is a point of contact with a curve of the family. If we substitute the coordinates x and y of any point of such a curve, and the slope p at the point, in (1) the equation is satisfied, since they are the x, y and p that belong to one of the curves given by (11). That is, the equation of the envelope is a solution of the differential equation. Being a single curve and not a particular solution, that is, not given by (11) for some particular value of c, the equation of the envelope is called the **singular solution.**

Singular solutions may be found by consideration of three conditions which they must satisfy.

(a) **The c discriminant.** Equation (11) gives, in general, two or more values of c for any point (x, y) as at A (Fig. 114). But at a point on the envelope, as B, certain values of c must coincide, giving fewer distinct values. The condition necessary that two or more values of c shall coincide is that the equations

$$F(x, y, c) = 0, \quad \frac{\partial F}{\partial c} = 0,$$

shall be satisfied for the same c. The elimination of c between these equations gives an equation

$$\theta_c(x, y) = 0,$$

and this function θ_c is called the c discriminant, and it must be made zero by the coordinates x, y of any point on the envelope. Therefore

the singular solution must be given by equating to zero some factor of the c discriminant.

(b) **The p discriminant.** Equation (1) gives, in general, two or more values of p for any point (x, y) as at A. But at a point on the envelope as at B, certain values of p must coincide. We may form the p discriminant by eliminating p between the equations

$$f(x, y, y) = 0, \quad \frac{\partial f}{\partial p} = 0.$$

Let the eliminant be

$$\theta_p(x, y) = 0$$

then θ_p is the p discriminant and the singular solution must be given by equating a factor of this p discriminant to zero.

FIG. 114

(c) **Test of a solution.** The c discriminant and the p discriminant must each contain the factor we seek, but each may contain other factors, giving loci which do not satisfy the differential equation. It is necessary, then to try the factors of at least one of the two discriminants, rejecting those which do not give solutions, that is, do not satisfy the test of a solution considered in Art. 62.

Of course any factor which gives a particular case under the general solution must be rejected.

As an example consider the equation

$$p^2 - (x + y)p + xy = 0.$$

The p discriminant is

$$(x + y)^2 - 4xy,$$

so that if there is a singular solution it must be $x - y = 0$, but putting $y = x$ and $p = 1$ in the equation above gives $1 - 2x + x^2 = 0$, which is not true in general, therefore there is no singular solution.

As a second example, take the differential equation

$$\left(\frac{d\rho}{d\theta}\right)^2 = 4\rho^2(\rho^2 - 1),$$

here $p = \frac{d\rho}{d\theta}$, and the two values coincide if $\rho = 0$ or $\rho = \pm 1$. Now $\rho = 0$ is a single point, the pole, but putting $\rho = \pm 1$ and $p = 0$ in the differential equation, we find the equation satisfied, therefore the equation of the unit circle is a singular solution. Again, writing the equation as

$$\frac{d\rho}{\rho\sqrt{\rho^2 - 1}} = \pm 2d\theta$$

we have

$$\sec^{-1}\rho = \pm 2\theta + c.$$

The ambiguous sign here is not significant, since $\sec(2\theta + c) = \sec(-2\theta + c_1)$ if $c_1 = -c$. The general solution then is

$$\rho = \sec(2\theta + c).$$

To find the c discriminant we take the partial derivative with respect to c,

$$0 = 2\sec(2\theta + c)\tan(2\theta + c),$$

and eliminating c we have

$$2\rho\sqrt{\rho^2 - 1} = 0,$$

giving again $\rho = 0$ or $\rho = \pm 1$. This gives again the unit circle, and it further shows that the pole, $\rho = 0$, is not a point on any curve of the family, therefore not a part of the envelope.

As a third example take

$$xp^2 - 2yp - x = 0.$$

The general solution is

$$c^2x^2 - 2cy - 1 = 0.$$

The p discriminant and the c discriminant are the same in this case, and each gives the equation

$$x^2 + y^2 = 0.$$

This represents a single real point, not on the curves in the general solution, but as an imaginary locus it satisfies the differential equation,

as it gives $p = -\frac{x}{y}$, and putting this in the differential equation reduces it to

$$\frac{x}{y^2}(x^2 + y^2) = 0.$$

EXERCISES LXII

1. Find the general solutions of the following differential equations by solving first for p:

(a) $p^2 + 3xp + 2x^2 = 0$; (c) $y^2 + p^2 = 1$;
(b) $xp^2 + (y - x^2)p = 0$; (d) $p^3 - p^2y - px^2 + x^2y = 0$.

2. Find the general solutions of the following by solving first for y:

(a) $p^2 + 2xp - y = 0$; (c) $y = px + p^3$;
(b) $4xp^2 + 2xp - y = 0$; (d) $p^2x - py + 4 = 0$.

3. In the following obtain the general solution by solving first for x.

(a) $x + py(2p^2 + 3) = 0$; (c) $p^2 + 3xp + 2x^2 = 0$;
(b) $y = px + p^3$; (d) $p^2 + 2xp - y = 0$.

Suggestion: In (d), upon reaching the differential equation in p and y regard p as independent variable and the equation becomes linear.

4. In solving 2(b) and 2(c), the differential equations in p and x contained certain factors free from $\frac{dp}{dx}$. Show that these factors give rise to singular solutions if we equate them to zero and eliminate p by use of the given equation.

5. Show that the same singular solution arises by the method of 3(b) as in 2(c).

6. The slope equation of tangents to the circle $x^2 + y^2 = 16$ is $y = mx + 4\sqrt{1 + m^2}$. Show that the differential equation of this family of lines is a Clairaut equation, and that its singular solution is the given circle.

7. The lines $y = mx - 4m - 2m^3$ are normal to the parabola $y^2 = 8x$. Find the differential equation and its singular solution.

8. Find the general and singular solutions of the differential equation

$$(y - px)^2 + 4p = 0.$$

9. Examine the following for singular solutions:

(a) $y^2 + p^2 = 1$;
(b) $p^2 + px - y = 0$;
(c) $\left(\frac{d\rho}{d\theta}\right)^2 + \rho^2 - \rho^4 = 0$.

CHAPTER XV

ORDINARY DIFFERENTIAL EQUATIONS, HIGHER ORDER

Art. 66. Homogeneous Linear Equations, Constant Coefficients

In this article we shall consider equations of the type

$$a_o \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \ldots + a_{n-1} \frac{dy}{dx} + a_n y = 0, \qquad (1)$$

in which the coefficients a_o, a_1, a_n are constants. The equation is **linear and homogeneous,** since y and its derivatives are involved linearly in each term. The equation being of the nth order, the general solution must contain n and not more than n arbitrary constants. Any function of x containing the required number of arbitrary constants and which, when put for y, causes (1) to be satisfied, will furnish the general solution. With this in mind, we may construct such a solution without actually integrating.

The fact that y and its successive derivatives are to have a linear relation suggests that y may have the form ce^{mx}, for such a function would differ from its derivatives only by a constant factor. We try then

$$y = ce^{mx}, \frac{dy}{dx} = cme^{mx}, \ldots \frac{d^n y}{dx^n} = cm^n e^{mx},$$

and substituting in (1) we find

$$ce^{mx}(a_o m^n + a_1 m^{n-1} + \ldots + a_{n-1} m + a_n) = 0. \qquad (2)$$

Dropping the factor ce^{mx}, which is not zero for any finite m, we have

$$a_o m^n + a_1 m^{n-1} + \ldots + a_{n-1} m + a_n = 0. \qquad (3)$$

This equation is known as the **auxiliary equation** to (1). There are n values of m which satisfy (3), and if we denote these values by m_1, m_2, ... , m_n, values of y which satisfy (1) are

$$c_1 e^{m_1 x}, c_2 e^{m_2 x}, \ldots c_n e^{mnx},$$

and if we write

$$y = c_1 e^{m_1 x} + c_2 e^{m_2 x} + \ldots + c_n e^{mnx} \qquad (4)$$

and substitute in (1) we shall find the sum of n expressions, each of the form (2), and all are zero separately. Therefore (4) is a solution of (1). **If the n values of m are all distinct (4) is the general solution,** but if some roots coincide a modification is necessary. Say $m_1 = m_2$, then the first two terms in the right hand member of (4) become a single term $(c_1 + c_2)e^{m_1 x}$, and the expression does not contain the required number of independent constants. If we try $y = xe^{mx}$, we have

$$\frac{dy}{dx} = (mx + 1)\, e^{mx}, \quad \frac{d^2y}{dx^2} = (m^2x + 2m)\, e^{mx} \quad . \quad . \quad . \quad .$$

$$\frac{d^ny}{dx^n} = (m^nx + nm^{n-1})\, e^{mx},$$

and substituting in (1) we have

$$xe^{mx}\,[a_0m^n + a_1m^{n-1} + \quad . \quad . \quad . \quad . + a_n] +$$

$$e^{mx}\,[a_0nm^{n-1} + a_1\,(n - 1)m^{n-2} + \quad . \quad . \quad . \quad . + a_{n-1}] = 0.$$

The first of these bracketed quantities vanishes if m satisfies (3) and the second, being the derivative with respect m of the first, must vanish if m is a double root of (3). Therefore, if $m_1 = m_2$, $y = xe^{m_1 x}$ is a solution of (1), and it may be shown in similar manner that if m_1 is an r-fold root of (3) then $y = x^s e^{m_1 x}$ is a solution if s is any integer less than r. **If, therefore, r values of m coincide, the corresponding terms of (4) may be multiplied respectively by** $1, x, x^2, \ldots . x^{r-1}$, **and the result is the general solution.**

One more modification of (4) is sometimes convenient. If among the roots of (3) there is a pair of complex roots, say $h \pm ki$, where $i = \sqrt{-1}$, then the corresponding terms in (4) may be written

$$c_1e^{(h+ki)x} + c_2e^{(h-ki)x} = e^{hx}\,(c_1e^{kix} + c_2e^{-kix})$$

$$= e^{hx}\,(c_1 \cos kx + c_1i \sin kx + c_2 \cos kx - c_2i \sin kx)$$

$$= e^{hx}\,(C_1 \cos kx + C_1 \sin kx), \qquad (5)$$

where $C_1 = c_1 + c_2$ and $C_2 = (c_1 - c_2)i$, or solving these for c_1 and c_2, $2c_1 = C_1 + C_2i$, $2c_2 = C_1 - C_2i$ (see Art. 51). We see, then, that **if m_1 and m_2 are conjugate complex roots, then the two terms of (4) take the real form (5) if c_1 and c_2 are given conjugate complex values.**

As an example consider

$$\frac{d^3y}{dx^3} - \frac{d^2y}{dx^2} - 8\,\frac{dy}{dx} + 12y = 0.$$

We write at once the auxiliary equation $m^3 - m^2 - 8m + 12 = 0$ and find the roots, $m = 2, 2, -3$. The general solution is

$$y = c_1e^{2x} + c_2xe^{2x} + c_3e^{-3x}.$$

Again take

$$\frac{d^4y}{dx^4} + \frac{d^3y}{dx^3} - 3\frac{d^2y}{dx^2} - 4\frac{dy}{dx} - 4y = 0.$$

Without writing the auxiliary equation we may find its roots, $2, -2, -\frac{1}{2} \pm \frac{1}{2}\sqrt{-3}$. The general solution is

$$y = c_1e^{2x} + c_2e^{-2x} + e^{-\frac{x}{2}}\left(c_3 \cos \frac{1}{2}\sqrt{3}\ x + c_4 \sin \frac{1}{2}\sqrt{3}\ x\right).$$

Two special cases are of extreme importance. The first is the **differential equation of harmonic motion,**

$$\frac{d^2y}{dx^2} + k^2y = 0. \tag{6}$$

Here the roots of the auxiliary equation are $\pm ki$, and the general solution is

$$y = c_1 \cos kx + c_2 \sin kx, \tag{7}$$

or, in a form sometimes more convenient

$$y = c \sin(kx + \phi). \tag{7'}$$

In this latter form c and ϕ are the arbitrary constants, and clearly $c^2 = c_1^2 + c_2^2$ and $c_1 = c \sin \phi$, $c_2 = c \cos \phi$. The function y is harmonic, the **amplitude** being c, the **period** $\frac{2\pi}{k}$ and the **phase angle** ϕ. It should be observed that the period of the function is controlled by the differential equation itself, while the amplitude and phase are arbitrary in the general solution, and may be determined to satisfy given initial conditions.

The second important case includes the first as a special type. The equation

$$\frac{d^2y}{dx^2} + 2a\frac{dy}{dx} + by = 0, \tag{8}$$

gives an auxiliary equation whose roots are $-a \pm \sqrt{a^2 - b}$. If b is greater than a^2 the general solution may be written

$$y = ce^{-ax} \sin (\sqrt{b - a^2}\, x + \phi). \tag{9}$$

If $a = 0$, this reduces to harmonic motion. If a is positive the factor e^{-ax} diminishes as x increases, giving oscillations diminishing in amplitude. In this case (9) is the equation and (8) the differential equation of **damped vibrations.** If a is negative (9) gives vibrations of increasing amplitude.

If in (8) we have a^2 greater than b, the solution may be written

$$y = e^{-ax} (c_1 e^{\sqrt{a^2-b}\,x} + c_2 e^{-\sqrt{a^2-b}\,x}). \tag{10}$$

If a is positive y approaches zero as x increases indefinitely, while if a is negative y increases with x. Another form for (10) is sometimes more convenient:

$$\begin{aligned} y &= e^{-ax} (C_1 \cosh \sqrt{a^2 - b}\, x + C_2 \sinh \sqrt{a^2 - b}\, x) \\ &= Ce^{-ax} \sinh (\sqrt{a^2 - b}\, x + \phi), \end{aligned} \tag{10'}$$

where $C_1 = c_1 + c_2$, $C_2 = c_1 - c_2$, $Ce^{\phi} = 2c_1$ and $Ce^{-\phi} = -2c_2$.

EXERCISES LXIII

1. Write the general solutions of the following:

(a) $\dfrac{d^2y}{dx^2} - 16y = 0$; (c) $\dfrac{d^3y}{dx^3} - 3\dfrac{d^2y}{dx^2} - \dfrac{dy}{dx} + 3y = 0$;

(b) $\dfrac{d^2y}{dx^2} + 5\dfrac{dy}{dx} + 6y = 0$; (d) $\dfrac{d^4y}{dx^4} - 6\dfrac{d^2y}{dx^2} + 8y = 0$.

2. Solve the following:

(a) $\dfrac{d^2y}{dx^2} = 0$; (c) $\dfrac{d^3y}{dx^3} - 6\dfrac{d^2y}{dx^2} + 9\dfrac{dy}{dx} = 0$;

(b) $\dfrac{d^4y}{dx^4} - 4\dfrac{d^2y}{dx^2} = 0$; (d) $4\dfrac{d^4y}{dx^4} - 4\dfrac{d^2y}{dx^2} + y = 0$.

3. Write in real form the general solutions of the following:

(a) $\dfrac{d^2y}{dx^2} + 9y = 0$; (c) $\dfrac{d^3y}{dx^3} - 2\dfrac{d^2y}{dx^2} + 5\dfrac{dy}{dx} = 0$;

(b) $\dfrac{d^4y}{dx^4} - 16y = 0$; (d) $\dfrac{d^5y}{dx^5} + 6\dfrac{d^3y}{dx^3} + 9\dfrac{dy}{dx} = 0$.

4. Solve the differential equation

$$\frac{d^2y}{dx^2} + 16y = 0,$$

subject to the conditions that $y = 0$ and $\frac{dy}{dx} = 5$ when $x = 0$.

5. Solve the equation

$$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 3y = 0$$

with the initial conditions $y = 2$ and $\frac{dy}{dx} = 1$ when $x = 0$.

6. Write the particular solution of

$$\frac{d^3y}{dx^3} + 4\frac{dy}{dx} = 0$$

for which $y = 0$ when $x = \frac{\pi}{4}$; $y = 1$ when $x = 0$, and $\frac{d^2y}{dx^2} = 0$ when $x = \frac{\pi}{8}$.

7. A spring deflects one inch under a 10 pound weight. If a weight of 40 pounds is attached to the spring, the force acting on the weight is $-120S$, where S is the distance in feet of the weight above its position of equilibrium, a deflection of 4 inches. Equating mass times acceleration to force, we have the differential equation

$$\frac{40}{g}\frac{d^2S}{dt^2} = -120S.$$

Show that the motion is harmonic, and determine the equation of motion assuming that motion starts when the spring is deflected 6 inches, that is, 2 inches below the position of equilibrium. What is the period of oscillation?

8. If in a situation such as 7 the motion is resisted by friction forces proportional to the velocity we should have a differential equation of the form

$$m\frac{d^2S}{dt^2} + n\frac{dS}{dt} = -kS.$$

Show that if n^2 is less than $4mk$ the motion consists of damped oscillations and find the frequency of oscillation.

9. If a round shaft of length L and diameter D is fixed at one end and twisted through an angle θ at the other the twisting moment is

$$M = \frac{\pi D^4 G \theta}{32L},$$

where G is the modulus of rigidity of the material of the shaft. If a weight W with radius of gyration k about the axis is attached to the rotating end of the shaft, and after angular displacement is left free under the twisting action of the shaft, we should have

$$\frac{Wk^2}{g}\frac{d^2\theta}{dt^2} = -\frac{\pi D^4 G \theta}{32L},$$

neglecting the inertia forces on the shaft itself. Find the period of the torsional vibrations.

10. If an electrical condenser of capacity K is charged with a quantity q the voltage is $\frac{q}{K}$. If the condenser is discharged through a circuit with self-induction L and resistance R we should have

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{K} = 0.$$

Find q as a function of t, taking $q = q_0$ and $\frac{dq}{dt} = 0$ when $t = 0$. Under what relations among L, R and K will the discharge consist of damped oscillations? Otherwise, what will be the character of the discharging current?

Art. 67. Non-homogeneous Linear Equations

Here we consider differential equations of the form

$$a_0\frac{d^ny}{dx^n} + a_1\frac{d^{n-1}y}{dx^{n-1}} + \ldots + a_ny = X, \tag{1}$$

in which the coefficients are again constants, and the quantity X on the right is a function of x only. If this function is identically zero, the equation reduces to the homogeneous case of the previous article.

We shall again make use of the auxiliary equation

$$a_0m^n + a_1m^{n-1} + \ldots + a_n = 0. \tag{2}$$

Using the roots of this as before, we may write

$$y = c_1e^{m_1x} + c_2e^{m_2x} + \ldots + c_ne^{m_nx} + f(x), \tag{3}$$

and this is a solution of (1) if on putting $y = f(x)$ the left hand member of (1) reduces to X, for the exponential terms written in (3) reduce the left hand member of (1) to zero, according to the previous article. If the roots of (2) are all distinct (3) is the general solution, containing the necessary number of arbitrary constants, otherwise the modifications used in the previous articles are effective. In case of complex roots of (2) the expression in terms of sine and cosine may be used as before.

The function f(x) is called a **particular integral** of (1), and the remainder of the general solution is called the **complementary function.**

To solve (1), then, we first replace X by zero, and solve the resulting homogeneous equation. Taking this solution as the complementary function we proceed to find a particular integral. In the three sections which follow we indicate three methods for finding the particular integral. The first is most convenient when it is applicable, and it is applicable in most practical cases. The other two methods are always applicable, and the last is important in preparation for the following chapter. In Section 4 we indicate a type of equation readily reducible to type (1).

1. **Trial particular integral.** If there exists a set of functions, $f_1(x)$, $f_2(x)$ . . . . $f_q(x)$, such that X and all of the successive derivatives of X may be represented as linear combinations of the form

$$b_1f_1(x) + b_2f_2(x) + \ldots\ldots b_qf_q(x),$$

then we may write

$$y = b_1f_1(x) + b_2f_2(x) + \ldots\ldots b_qf_q(x) \tag{4}$$

and determine the b's so that this is a particular integral by substituting in (1) and equating coefficients of like terms on the right and left. Such a set of functions, if it exists, may be discovered by writing the terms of X and their successive derivatives, until on further differentiation no new terms arise.

Quite evidently if a term in (4) is a term in the complementary function its coefficient cannot be determined, since it becomes an arbitrary c of the general solution. On this account **an exception must be made if a term in X is of the form kx^pe^{mx}, where ce^{mx} is a term of the complementary function. If m is an r-fold root of the auxiliary equation we replace this term of X by $kx^{p+r}e^{mx}$ in making up the trial integral, and omit from (4) such terms as are included in the complementary function,** already known.

A general demonstration would require too much space to be profitable

here, but the method is justified by its results, and may be illustrated by examples. Consider

$$\frac{d^4y}{dx^4} - 9\frac{d^2y}{dx^2} = x^3 - e^{3x}.$$

The roots of the auxiliary equation are 0, 0, ± 3, and the complementary function is

$$c_1 + c_2x + c_3e^{3x} + c_4e^{-3x}.$$

We may seek the particular integral in two parts, the first part to yield x^3, the second part to yield $-e^{3x}$. For the first part we must take account of the exception, and in the notation above we have $k = 1$, $p = 3$, $m = 0$, $r = 2$. We replace x^3, then, by x^5, and the trial integral is

$$y = b_1x^5 + b_2x^4 + b_3x^3 + b_4x^2.$$

The required derivatives are

$$\frac{d^2y}{dx^2} = 20b_1x^3 + 12b_2x^2 + 6b_3x + 2b_4,$$

$$\frac{d^4y}{dx^4} = 120b_1x + 24b_2.$$

Substituting in the differential equation and equating coefficients we find

$$-180b_1 = 1,\ 108b_2 = 0,\ 120b_1 - 54b_3 = 0,\ 24b_2 - 18b_4 = 0,$$

so that the part of the particular integral found is

$$y = \frac{-x^5}{180} - \frac{x^3}{81}.$$

Now for the second part of the particular integral we must again take account of the exception, and replace e^{3x} by xe^{3x}, and we take for a trial integral

$$y = bx\,e^{3x}.$$

The required derivatives are

$$\frac{d^2y}{dx^2} = 9bxe^{3x} + 6be^{3x},\ \frac{d^4y}{dx^4} = 81bxe^{3x} + 108be^{3x}$$

And substituting and equating coefficients we find $54b = -1$, and the general solution is

$$y = c_1 + c_2x + c_3e^{3x} + c_4e^{-3x} - \frac{x^5}{180} - \frac{x^3}{81} - \frac{1}{54}xe^{3x}.$$

As a second example take

$$\frac{d^3y}{dx^3} - 2\frac{d^2y}{dx^2} + 4\frac{dy}{dx} - 8y = x \sin 2x.$$

Here the complementary function may be written

$$c_1e^{2x} + c_2 \sin 2x + c_3 \cos 2x,$$

and we find that the right hand side of the differential equation is a term in the complementary function multiplied by x, and that the corresponding root of the auxiliary equation is not multiple. Accordingly we use in the trial integral the function $x^2 \sin 2x$ and terms arising by differentiation but not in the complementary function. The trial integral is

$$y = b_1x^2 \sin 2x + b_2x^2 \cos 2x + b_3x \sin 2x + b_4x \cos 2x,$$

and forming the derivatives and substituting in the differential equation we may find the values of the b's which make this a particular integral.

2. **Variation of parameters.** Having formed the complementary function, instead of equation (3) above, we may write the equation

$$y = c_1e^{m_1x} + c_2e^{m_2x} + \ldots + c_ne^{m_nx}, \qquad (5)$$

making the proper modifications for multiple and complex roots then instead of regarding the c's as constants we may regard them as functions of x and inquire what functions they may be in order that (5) shall be a solution of (1). The method is known as **variation of parameters.**

On substituting y and its derivatives in (1) we find one condition that must be fulfilled by the n arbitrary functions, $c_1, c_2 \ldots c_n$, and we may choose $n - 1$ other conditions according to our own convenience, the n equations thus written being sufficient to determine the functions. In practice the conditions are so chosen that the n equations are linear in the first derivatives of the c's. The use of the method may be illustrated by solving the last example above,

$$\frac{d^3y}{dx^3} - 2\frac{d^2y}{dx^2} + 4\frac{dy}{dx} - 8y = x \sin 2x.$$

We write

$$y = c_1e^{2x} + c_2 \sin 2x + c_3 \cos 2x$$

$$\frac{dy}{dx} = 2c_1e^{2x} + 2c_2 \cos 2x - 2c_3 \sin 2x + e^{2x}\frac{dc_1}{dx} + \sin 2x \frac{dc_2}{dx} + \cos 2x \frac{dc_3}{dx}.$$

Now, as one of the two conditions which we may choose at will, we set the sum of these last three terms equal to zero, and obtain for the second derivative

$$\frac{d^2y}{dx^2} = 4c_1e^{2x} - 4c_2 \sin 2x - 4c_3 \cos 2x + 2e^{2x}\frac{dc_1}{dx} + 2 \cos 2x \frac{dc_2}{dx} - 2 \sin 2x \frac{dc_3}{dx},$$

As a second condition we set the sum of the last three terms here equal to zero, then the third derivative becomes

$$\frac{d^3y}{dx^3} = 8c_1e^{2x} - 8c_2 \cos 2x + 8c_3 \sin 2x + 4e^{2x}\frac{dc_1}{dx} - 4 \sin 2x \frac{dc_2}{dx} - 4 \cos 2x \frac{dc_3}{dx}.$$

Now, if we substitute y and its three derivatives in the differential equation, remembering that the first three terms of each derivative are the same as derivatives of the complementary function, we see that the equation is satisfied if the sum of the last three terms written is the function x sin 2x. We have, then, the three conditions

$$e^{2x}\frac{dc_1}{dx} + \sin 2x \frac{dc_2}{dx} + \cos 2x \frac{dc_3}{dx} = 0,$$

$$2e^{2x}\frac{dc_1}{dx} + 2 \cos 2x \frac{dc_2}{dx} - 2 \sin 2x \frac{dc_3}{dx} = 0,$$

$$4e^{2x}\frac{dc_1}{dx} - 4 \sin 2x \frac{dc_2}{dx} - 4 \cos 2x \frac{dc_3}{dx} = x \sin 2x.$$

We observe that these equations may be written without carrying out all the steps, since the coefficients of the derivatives of the c's are the

coefficients of the c's in the complementary function and their successive derivatives. Solving the three linear algebraic equations we find

$$\frac{dc_1}{dx} = \frac{x}{8} e^{-2x} \sin 2x,$$

$$\frac{dc_2}{dx} = -\frac{x}{8} \sin 2x (\sin 2x + \cos 2x),$$

$$\frac{dc_3}{dx} = \frac{x}{8} \sin 2x (\sin 2x - \cos 2x).$$

We may integrate to find the c's, omitting the constants of integration, since we are seeking only a particular integral and adding the constants would result only in producing again the complementary function which is already known. We have

$$c_1 = -\frac{1}{64} e^{-2x} (\cos 2x + 2x \sin 2x + 2x \cos 2x),$$

$$c_2 = -\frac{x^2}{32} + \frac{1}{256} (-\sin 4x + \cos 4x + 4x \sin 4x + 4x \cos 4x),$$

$$c_2 = \frac{x^2}{32} + \frac{1}{256} (-\sin 4x - \cos 4x - 4x \sin 4x + 4x \cos 4x),$$

and substituting these in the expression for y, reducing and dropping terms that repeat terms of the complementary function, we have for the general solution

$$y = c_1 e^{2x} + (c_2 - 3x - 2x^2) \sin 2x + (c_3 - x + 2x^2) \cos 2x.$$

3. **Use of symbolic operator.** The notation $D_x y$ is very commonly used to denote the derivative of y with respect to x, and in case no ambiguity is occasioned the independent variable as a subscript is omitted, giving the notation Dy. Similarly the successive derivatives of y may be denoted by D^2y, D^3y, etc. The differential equation (1) may be written

$$a_0 D^n y + a_1 D^{n-1} y + \ldots + a_n y = X. \tag{6}$$

Now it is convenient to write this equation in the form

$$[a_0 D^n + a_1 D^{n-1} + \ldots a_n] y = X \tag{7}$$

and regard the expression in brackets as the symbol of an operation, just as D alone is the symbol of an operation. The operation on y

indicated in (7) is, of course, equivalent to performing the operations indicated in (6) or (1) and adding the results. The expression in brackets is called a **symbolic operator.**

If we regard the symbol of this operator as an algebraic expression we may factor it and write the equation

$$(D - m_1)(D - m_2) \ . \ . \ . \ . \ (D - m_n)y = X, \tag{8}$$

where the m's are the roots of the auxiliary equation. Now each of these factors has a definite meaning as a symbolic operator, as

$$(D - m)y = Dy - my = \frac{dy}{dx} - my.$$

Thus the symbol which indicates multiplication in algebra is used here to indicate an operation of another kind. In algebra (8) would indicate multiplication by several factors in succession, and if we here interpret the symbol as indicating several operations in succession the result is the same as that in (7) or (6). It will suffice here to show this in case n = 2.

$$(D - m_1)(D - m_2)y = (D - m_1)(Dy - m_2y)$$

$$= D(Dy - m_2y) - m_1(Dy - m_2y)$$

$$= D^2y - m_2Dy - m_1Dy + m_1m_2y = [D^2 - (m_1 + m_2)D + m_1m_2]y.$$

Knowing these properties of the symbolic operator, we may write

$$\begin{aligned} (D - m_n)y &= u_1, \\ (D - m_{n-1})u_1 &= u_2, \\ &\cdots\cdots\cdots\cdots \\ (D - m_1)u_{n-1} &= X, \end{aligned} \tag{9}$$

where the u's are auxiliary variables, functions of x, and we may find u_{n-1} by solving the last of these equations, then successively solving the other equations of (9) we finally obtain y. The equations are all linear of the first order and are easily solved. The constants of integration may be omitted, since they merely introduce again the terms of the complementary function.

An example will make the method clear. Take

$$D^2y - 4Dy + 3y = \frac{1}{x}.$$

This may be written, corresponding to (8),

$$(D - 3)(D - 1)y = \frac{1}{x}.$$

Equations (9) become

$$(D - 3)u = \frac{1}{x}, \quad (D - 1)y = u.$$

We have then, using the method of Art. 64,

$$ue^{-3x} = \int \frac{1}{x} e^{-3x}\, dx,$$

and using this value of u the second differential equation becomes

$$Dy - y = e^{3x} \int \frac{1}{x} e^{-3x}\, dx,$$

and this gives for y the equation

$$ye^{-x} = \int \left[e^{2x} \int \frac{1}{x} e^{-3x}\, dx \right] dx.$$

The integration indicated can be accomplished only by the use of infinite series. If we apply first integration by parts we find

$$ye^{-x} = \frac{1}{2} e^{2x} \int \frac{1}{x} e^{-3x}\, dx - \frac{1}{2} \int \frac{1}{x} e^{-x}\, dx,$$

then expanding the exponential factors of the integrands in series, dividing by x, and integrating term by term, we arrive at the general solution.

$$y = \frac{1}{2} e^{3x} \left[c_1 + \log x - 3x + \frac{9}{4} x^2 - \frac{9}{6} x^3 + \ldots \right]$$
$$- \frac{1}{2} e^{x} \left[c_2 + \log x - x + \frac{x^2}{4} - \frac{x^3}{18} + \ldots \right].$$

4. **Legendre's linear equation.** An equation of the type

$$k_o(a + bx)^n \frac{d^n y}{dx^n} + \ldots + k_{n-1}(a + bx) \frac{dy}{dx} + k_n = X \quad (10)$$

may be reduced to a linear equation with constant coefficients by the substitution

$$a + bx = e^z.$$

We find for the first and second derivatives,

$$\frac{dy}{dx} = \frac{dy}{dz} \cdot \frac{dz}{dx} = \frac{dy}{dz} \cdot \frac{b}{a + bx},$$

$$\frac{d^2y}{dx^2} = \frac{d^2y}{dz^2}\left(\frac{dz}{dx}\right)^2 + \frac{dy}{dz}\frac{d^2z}{dx^2} = \left(\frac{d^2y}{dz^2} - \frac{dy}{dz}\right)\frac{b^2}{(a + bx)^2},$$

and so for the nth derivative of y with respect to x we find an expression which when multiplied by $(a + bx)^n$ becomes linear in the derivatives of y with respect to z, with constant coefficients. The function X on the right becomes a function of z when we put $x = \frac{1}{b}(e^z - a)$. We obtain, then, a differential equation of the form previously considered, with z as independent variable. When the solution is found we put $z = \log(a + bx)$ to restore the original variables.

The equation is known as **Legendre's linear equation,** while the name of Cauchy is associated with the special case when $a = 0$ and $b = 1$.

EXERCISES LXIV

1. Solve the differential equations

(a) $\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 6y = 12 + x;$

(b) $\frac{d^4y}{dx^2} - 4\frac{d^2y}{dx^2} = \cos x;$

(c) $\frac{d^3y}{dx^3} - 6\frac{d^2y}{dx^2} + 9\frac{dy}{dx} = e^{-x}.$

2. Find solutions for

(a) $\frac{d^4y}{dx^4} - 4\frac{d^2y}{dx^2} = 1 + e^{2x};$

(b) $\frac{d^2y}{dx^2} + 4y = \sin 2x;$

(c) $\frac{d^3y}{dx^3} - 2\frac{d^2y}{dx^2} + 5\frac{dy}{dx} = x^2 - \cos 4x.$

3. Solve by variation of parameters:

(a) $\dfrac{d^2y}{dx^2} + y = \tan x;$

(b) $\dfrac{d^2y}{dx^2} - 4\dfrac{dy}{dx} + 3y = \dfrac{1}{x};$

(c) $\dfrac{d^3y}{dx^3} - 3\dfrac{d^2y}{dx^2} - \dfrac{dy}{dx} + 3y = 8e^{3x}\cos x.$

4. Solve the following:

(a) $(D - 1)^2y = e^x \sec^2 x;$

(b) $(D - 2)(D + 1)y = e^{-x}(1 - 3x).$

5. (a) Solve the Cauchy linear equation

$$x^2\frac{d^2y}{dx^2} + 5x\frac{dy}{dx} + 4y = x.$$

(b) Solve the Legendre linear equation

$$(2x - 1)^2\frac{d^2y}{dx^2} + 2(2x - 1)\frac{dy}{dx} - 36y = 1.$$

6. If the body in example 7, Art. 66, is acted upon by an impressed vertical force that is periodic, in addition to its weight and the pull of the spring, the differential equation takes the form

$$\frac{40}{g}\frac{d^2s}{dt^2} + 120s = F\sin\omega t,$$

where the constant F is the maximum value of the impressed force and $\dfrac{2\pi}{\omega}$ is its period. Discuss the motion: (a) when $\omega = \sqrt{3g}$ (b) when ω has any other value.

7. If in example 8, Art. 66, we add a periodic force $F\sin\omega t$, assuming that n^2 is less than $4mk$, show that the motion rapidly approaches a steady vibration of the type

$$s = k\sin(\omega t + \phi),$$

and find for what value of ω the amplitude will be a maximum.

8. If a column AB with rounded ends (Fig. 115) is supported laterally at B and both vertically and laterally at A, and sustains a load P

acting at a distance d from the axis of the column, the bending moment on the column at any distance x from a is approximately

$$EI\frac{d^2y}{dx^2} = -Py - \frac{Pd}{L}x,$$

where E is the modulus of elasticity of the material, I is moment of inertia of the section area, L is the length of the column and y is the lateral displacement at distance x. Find the elastic curve, noting that $y = 0$ when $x = 0$ and when $x = L$.

9. In example 10, Art. 66, add an impressed electromotive force $E \sin \omega t$, and, in the solution of the differential equation discriminate between "transient" and "permanent" phenomena.

Fig. 115

Art. 68. Solution by Reduction of Order

In this article we shall consider several devices which, under appropriate conditions, may have the effect of reducing the order of the differential equation. The process of solving a differential equation of given order may thus be accomplished by the solution of a chain of equations of lower order, usually of the first order.

1. **When y is absent.** If the differential equation does not contain y explicitly, then we may reduce the order by a change of the dependent variable. Let the order of the equation be n, and let the lowest ordered derivative present be q, then putting

$$u = \frac{d^q y}{dx^q}, \frac{du}{dx} = \frac{d^{q+1}y}{dx^{q+1}}, \ldots\ldots, \frac{d^{n-q}u}{dx^{n-q}} = \frac{d^n y}{dx^n},$$

we reduce the order of the equation to $n - q$.

For example, take

$$x\frac{d^2y}{dx^2} - 2\frac{dy}{dx} = x^3 \sin x.$$

Putting $u = \dfrac{dy}{dx}$ we have the linear equation of the first order

$$\frac{du}{dx} - \frac{2u}{x} = x^2 \sin x.$$

The integrating factor is $e^{\int -\frac{2dx}{x}} = e^{\log \frac{1}{x^2}} = \dfrac{1}{x^2}$, therefore

$$\frac{u}{x^2} = \int \sin x \, dx = -\cos x + c,$$

and

$$u = \frac{dy}{dx} = -x^2 \cos x + cx^2,$$

and the general solution is

$$y = -x^2 \sin x - 2x \cos x + 2 \sin x + c_1x^3 + c_2.$$

2. **When x is absent.** If x does not appear explicitly in the differential equation, then the order of the equation is made one lower by writing

$$\frac{dy}{dx} = u,$$

$$\frac{d^2y}{dx^2} = \frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx} = u\frac{du}{dy},$$

$$\frac{d^3y}{dx^3} = \frac{du}{dx}\frac{du}{dy} + u\frac{d^2u}{dy^2}\frac{dy}{dx} = u\left(\frac{du}{dy}\right)^2 + u^2\frac{d^2u}{dy^2},$$

and so on, each successive derivative of y with respect to x being expressible in terms of u and derivatives of u with respect to y, always of lower order. Substituting in the differential equation we find a new differential equation of lower order, with y as independent variable and u as dependent variable.

Consider the example

$$y\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2 = 1.$$

We have

$$yu\frac{du}{dy} + u^2 = 1,$$

and separating the variables and integrating we find

$$(u^2 - 1)y^2 = c$$

and, remembering that $u = \dfrac{dy}{dx}$, we obtain the general solution in the form

$$y^2 = x^2 + c_1x + c_2.$$

3. **Exact differential equations.** In Art. 64 we considered exact differential equations of the first order. A differential equation of any order is said to be exact if it is the direct result of taking the derivative of an equation of lower order. No general criterion enables us to test directly for exactness, but we may proceed by inspection to investigate a given equation and find its solution in case it is exact.

Let us consider from this point of view the equation just solved above,

$$y\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2 - 1 = 0.$$

We see that the first term would arise on differentiation of $y\dfrac{dy}{dx}$, and differentiating we have

$$\frac{d}{dx}\left(y\frac{dy}{dx}\right) = y\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2,$$

and subtracting this from the given expression we see that we still require something whose derivative is -1, and that is obviously $-x + c$, so that the given equation is the exact derivative of

$$y\frac{dy}{dx} - x + c = 0.$$

And now integrating this we obtain the general solution.

As another illustration consider

$$x\frac{d^2y}{dx^2} - y\frac{dy}{dx} + 4x = 0.$$

As a first step we write

$$\frac{d}{dx}\left(x\frac{dy}{dx}\right) = x\frac{d^2y}{dx^2} + \frac{dy}{dx},$$

and by subtraction we see that it is necessary to find a quantity whose derivative is

$$-y\frac{dy}{dx} - \frac{dy}{dx} + 4x,$$

and the quantity is evidently

$$-\frac{1}{2}y^2 - y + 2x^2 + c,$$

so that the given equation is the exact derivative of

$$x\frac{dy}{dx} - \frac{1}{2}y^2 - y + 2x^2 + c = 0.$$

We thus reduced the problem of solving the given equation to that of solving one of lower order. The new differential equation of lower order that arises in this manner is called the **first integral** of the given equation. In some practical problems this first integral has a significance as important as that of the complete solution.

EXERCISES LXV

1. Solve the following equations:

(a) $\dfrac{d^2y}{dx^2} = \sqrt{1 + \left(\dfrac{dy}{dx}\right)^2}$;

(b) $x\dfrac{d^2y}{dx^2} + \dfrac{dy}{dx} = x^2$;

(c) $\dfrac{d^2y}{dx^2} + x\dfrac{dy}{dx} = x$.

2. Solve: (a) $2\dfrac{d^2y}{dx^2} = e^y$;

(b) $\dfrac{d^2y}{dx^2} + k^2y = 0$; (c) $\dfrac{d^2y}{dx^2} + \dfrac{a^2}{y^2} = 0$.

3. Solve the differential equations:

(a) $(x^3 - x)\dfrac{d^2y}{dx^2} + (5x^2 - 2)\dfrac{dy}{dx} + 4xy = 0;$

(b) $(5x^2 - 2)\dfrac{d^2y}{dx^2} + 14x\dfrac{dy}{dx} + 4y = 0;$

(c) $(x^3 - x)\dfrac{d^3y}{dx^3} + (8x^2 - 3)\dfrac{d^2y}{dx^2} + 14x\dfrac{dy}{dx} + 4y = 0.$

4. Show that the only curves with constant curvature are circles by solving the differential equation

$$\frac{d^2y}{dx^2} = k\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{\frac{3}{2}}.$$

CHAPTER XVI

ORDINARY DIFFERENTIAL EQUATIONS, SEVERAL VARIABLES

Art. 69. Systems of Simultaneous Equations

Here we shall consider systems of differential equations involving a single independent variable, t, and several dependent variables, x, y, z, etc., functions of t, together with one or more derivatives of these variables with respect to t. The solution of such a system consists in finding expressions for x, y, z, etc., in terms of t, which, when substituted in the given equations, reduce each equation to an identity in t. For the general solution the expressions for x, y, z, etc., must contain a certain number of arbitrary constants, the number depending on the order of the derivatives present in the given system of equations.

The given systems must contain as many independent equations as there are dependent variables, and we may form additional equations by differentiating these with respect to t as many times as desired. From this enlarged system we may, by algebraic processes, eliminate all dependent variables and their derivatives excepting one, say x, and its derivatives, obtaining a single differential equation involving x and t and derivatives of x with respect to t. This may be solved, giving x in terms of t and a certain number of arbitrary constants. The process may be repeated for y, z, etc. After any variable is found as a function of t the system may be reduced, if desired, by substituting this value in the given equations before solving for the remaining variables. The arbitrary constants that appear in the expressions for x, y, z, etc., may not be independent, but relations among them are found by substituting in the original equations, which must be identities in t.

For example, consider the two equations

$$\left(\frac{d^2x}{dt^2}\right)^2 = \left(\frac{d^2y}{dt^2}\right)^2 + 1, \tag{1}$$

$$\frac{dy}{dt} - x = t + 4. \tag{2}$$

By differentiating (2) we have

$$\frac{d^2y}{dt^2} - \frac{dx}{dt} = 1. \tag{3}$$

From (1) and (3) we obtain a differential equation in x and t only,

$$\left(\frac{dx^2}{dt^2}\right)^2 = \left(\frac{dx}{dt} + 1\right)^2 + 1, \tag{4}$$

and the general solution of this is

$$x = \cosh (t + C_1) - t + C_2. \tag{5}$$

If we substitute this in (2) and integrate we find

$$y = \sinh (t + C_1) + C_2 t + 4t + C_3. \tag{6}$$

Equations (5) and (6) together give the general solution of the simultaneous equations (1) and (2).

Instead of finding y from (2) after reducing by means of (5) we might have differentiated (3), obtaining

$$\frac{d^3y}{dt^2} - \frac{d^2x}{dt^2} = 0, \tag{7}$$

and this with (1), leads to the equation in y and t only,

$$\left(\frac{d^3y}{dt^3}\right)^2 = \left(\frac{d^2y}{dt^2}\right)^2 + 1. \tag{8}$$

The general solution of this is

$$y = \sinh (t + b_1) + b_2 t + b_3. \tag{9}$$

Equations (5) and (9) give the general solution, but the constants are not independent. If we substitute (5) and (9) in (2) we find that $b_1 = C_1$, $b_2 = C_2 + 4$, while b_3 is arbitrary.

It may be observed that elimination of y and its derivatives from (1) and (2) leads definitely to (4), and therefore to (5), and that elimination of x and its derivatives leads definitely to (8) and (9). Just so, associated with any system of n differential equations in n dependent variables there will be n definite equations of higher order, each involving only one dependent variable. But the general solutions of these (n) separate equations do not give the solution of the original n simultaneous equations until the arbitrary constants arising are so related as to satisfy the given equations. It is not possible, in general, to return from the separate equations, as (4) and (8), to the original system (1) and (2), since many different systems, other than (1) and (2), might yield the same (4) and (8). The differences in the general

solutions lie in the different relations that might exist between the constants of (5) and (9).

In general, we may associate with any system of differential equations a system of equations of the first order. This is accomplished by putting an auxiliary variable for each derivative present, excepting the highest present in each variable. Thus, from (1) and (2) we may have

$$\frac{dx}{dt} = u,$$

$$\frac{dy}{dt} = v,$$

$$\left(\frac{du}{dt}\right)^2 = \left(\frac{dv}{dt}\right)^2 + 1,$$

$$v - x = t + 4.$$

The fourth of these is not a differential equation, and using this to eliminate v we have

$$\frac{dx}{dt} = u,$$

$$\frac{dy}{dt} = x + t + 4,$$

$$\frac{du}{dt} = [(u + 1)^2 + 1]^{\frac{1}{2}}. \tag{10}$$

This system of three equations leads as definitely to (4) and (8) as does the original system (1) and (2). Moreover, from the system (10) we may find solution (5) and (6) at once. The third equation involves u and t only, giving

$$u = \sinh (t + C_1),$$

then the first gives (5) and the second gives (6).

The difficulties met in the solution of systems of differential equations usually pertain to the algebraic process of elimination and to the process of solving the single differential equations after the elimination is accomplished. Various special devices may be employed for systems of particular types. In the two sections below we indicate the most important cases.

1. **Systems of linear equations with constant coefficients.** First let us consider an example. Given the two equations

$$\frac{d^2y}{dt^2} - \frac{dx}{dt} - 2y = 2t, \tag{11}$$

$$4\frac{dx}{dt} + \frac{dy}{dt} - 3x = 0, \tag{12}$$

we may, by differentiating (11) once and (12) twice, obtain the additional equations

$$\frac{d^3y}{dt^3} - \frac{d^2x}{dt^2} - 2\frac{dy}{dt} = 2, \tag{13}$$

$$4\frac{d^2x}{dt^2} + \frac{d^2y}{dt^2} - 3\frac{dx}{dt} = 0, \tag{14}$$

$$4\frac{d^3x}{dt^3} + \frac{d^3y}{dt^3} - 3\frac{d^2x}{dt^2} = 0. \tag{15}$$

From these five equations we may eliminate y, and its three successive derivatives, reaching a single equation in x and t,

$$4\frac{d^3x}{dt^3} - 2\frac{d^2x}{dt^2} - 8\frac{dx}{dt} + 6x = -2. \tag{16}$$

Now, using the notion of differential operator introduced in Art. 66, we may let the symbol D denote differentiation with respect to t, and the equations may be written

$$D^2y - Dx - 2y = 2t, \tag{11'}$$

$$4Dx + Dy - 3x = 0, \tag{12'}$$

$$D^3y - D^2x - 2Dy = D(2t), \tag{13'}$$

$$4D^2x + D^2y - 3Dx = 0, \tag{14'}$$

$$4D^3x + D^3y - 3D^2x = 0. \tag{15'}$$

The last three equations might have been derived algebraically, using D as a multiplier, and since the elimination is algebraic, the whole process is equivalent to an algebraic process. Accordingly, we may substitute a more direct algebraic method. Writing (11) and (12) in the form

$$Dx + (D^2 - 2)y = 2t,$$

$$(4D - 3)x + Dy = 0,$$

and solving for x by determinants, regarding x and y as unknowns,

$$\begin{vmatrix} D & D^2 - 2 \\ 4D - 3 & D \end{vmatrix} x = \begin{vmatrix} 2t & D^2 - 2 \\ 0 & D \end{vmatrix}.$$

Expanding these determinants we have

$$(-4D^3 + 2D^2 + 8D - 6)x = D(2t) = 2, \tag{16'}$$

which is equivalent to (16).

Similarly, solving for y we have

$$\begin{vmatrix} D & D^2 - 2 \\ 4D - 3 & D \end{vmatrix} y = \begin{vmatrix} D & 2t \\ 4D - 3 & 0 \end{vmatrix}$$

or

$$(-4D^3 + 2D^2 + 8D - 6)y = -(4D - 3)(2t) = 6t - 8. \tag{17}$$

From (16) we have

$$x = (C_1 + C_2t)e^t + C_3e^{-\frac{3}{2}t} - \frac{1}{3},$$

and from (17)

$$y = (C_4 + C_5t)e^t + C_6e^{-\frac{3}{2}t} - t.$$

Substituting these values in (12) we find that

$$C_4 = -C_1 - 3C_2,\ C_5 = -C_2,\ C_6 = -6C_3.$$

Again, consider the system

$$\begin{aligned} \frac{dz}{dt} + 4x &= \sin t, \\ \frac{dy}{dt} + \frac{dx}{dt} &= 2\cos t, \\ \frac{dy}{dt} + z &= 0. \end{aligned} \tag{18}$$

This may be written, by use of the differential operator, in the form

$$\begin{aligned} 4x + 0y + Dz &= \sin t, \\ Dx + Dy + 0z &= 2\cos t, \\ 0x + Dy + z &= 0. \end{aligned}$$

Solving for x we have

$$\begin{vmatrix} 4 & 0 & D \\ D & D & 0 \\ 0 & D & 1 \end{vmatrix} x = \begin{vmatrix} \sin t & 0 & D \\ 2\cos t & D & 0 \\ 0 & D & 1 \end{vmatrix},$$

from which we reach the equation

$$(D^3 + 4D)\, x = -\cos t.$$

The general solution of this is

$$x = C_1 + C_2 \sin 2t + C_3 \cos 2t - \frac{1}{3} \sin t,$$

and substituting this in the first and second of equations (18) we have for z and y,

$$z = -4C_1 t + 2C_2 \cos 2t - 2C_3 \sin 2t - \frac{7}{3} \cos t + C_4,$$

$$y = -C_2 \sin 2t - C_3 \cos 2t + \frac{7}{3} \sin t + C_5.$$

If we substitute these values in the third equation of (18) we find $C_1 = 0$ and $C_4 = 0$, so that the general solution of the system (18) is

$$x = C_2 \sin 2t + C_3 \cos 2t - \frac{1}{3} \sin t,$$

$$y = -C_2 \sin 2t - C_3 \cos 2t + \frac{7}{3} \sin t + C_5,$$

$$z = -2C_3 \sin 2t + 2C_2 \cos 2t - \frac{7}{3} \cos t.$$

From consideration of the examples treated here the method for dealing with any system of linear equations with constant coefficients should be sufficiently evident.

2. **Systems of equations of the first order.** If every equation of a system is of the first order we may suppose the system solved for the derivatives, giving, in case of three dependent variables, a system of the form

$$\frac{dx}{dt} = P, \frac{dy}{dt} = Q, \frac{dz}{dt} = R. \tag{19}$$

where P, Q and R may be functions of x, y, z and t. An equivalent form for the system is

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R} = \frac{dt}{1}. \tag{20}$$

Here the denominators may be multiplied or divided by any quantity desired. The form (20) may be extended to any number of variables, and offers many advantages in the process of elimination. If any of the equations

$$\frac{dx}{P} = \frac{dy}{Q}, \frac{dx}{P} = \frac{dz}{R}, \frac{dy}{Q} = \frac{dz}{R},$$

etc., should afford a differential equation in the two variables whose differentials appear, then its solution affords one relation, with an arbitrary constant, by which the system may be reduced. The general solution is known when the number of relations so found is equal to the number of dependent variables, for these relations permit each dependent variable to be expressed as a function of t.

The system (10) in the previous section is an example of form (19). Reduced to form (20) it gives

$$\frac{du}{[(u+1)^2+1]^{\frac{1}{2}}} = \frac{dy}{x+t+4} = \frac{dx}{u} = \frac{dt}{1}.$$

The first and last or first and third of these give equations that may be solved, then by substitution the solution may be completed.

Another method for finding relations among the variables in (20) is based on the identity

$$\frac{Xdx + Ydy + Zdz + Tdt}{XP + YQ + ZR + T} = \frac{dt}{1}. \tag{21}$$

Here X, Y, Z and T are any functions of the variables, and the method consists in choosing them so that the resulting differential equation may be solved. Of course the right member of (21) may be replaced by any other member of (20) or by another expression of the form of the left member of (21).

To illustrate, consider the system

$$\frac{dx}{y+t} = \frac{dy}{x-t} = \frac{dt}{x+y}.$$

Using (21), but taking the functions X and Y as each equal to unity, we may write

$$\frac{dx + dy}{x + y} = \frac{du}{u} = \frac{dt}{u},$$

where $u = x + y$. This gives $x + y = t + C_1$. Now substituting the value of x in terms of y and t we have

$$\frac{dy}{C_1 - y} = \frac{dt}{t + C_1},$$

from which we find $(C_1 - y)(t + C_1) = C_2$. We may now write the general solution in the form

$$x = t + \frac{C_2}{t + C_1}, \quad y = C_1 - \frac{C_2}{t + C_1}.$$

Further uses of (21) in solving systems of differential equations of the first order will be discussed in the next article.

EXERCISES LXVI

1. Solve the system of differential equations for motion of an unresisted projectile,

$$\frac{d^2x}{dt^2} = 0, \quad \frac{d^2y}{dt^2} = -g.$$

2. In a certain plane conservative force field the potential function is $\frac{1}{2}k^2m(x^2 + y^2)$, where m is the mass of the body acted upon. The equations of motion of the body are, therefore,

$$\frac{d^2y}{dt^2} = -k^2y, \quad \frac{d^2x}{dt^2} = -k^2x.$$

Show that the path of the body is a conic section.

3. Solve the system

$$\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 = 1,$$

$$t\frac{dy}{dt} + \frac{dx}{dt} = y.$$

4. Solve the system

$$\frac{dx}{dt} + \frac{dy}{dt} - 2x + y = 3e^t,$$

$$6\frac{dx}{dt} + \frac{dy}{dt} + 4x + y = 8e^{2t}.$$

5. Solve the system

$$\frac{dx}{dt} + x + 2y = \sin t,$$

$$9\frac{dy}{dt} - 9y - 4x = t.$$

6. Solve the system

$$\frac{d^2y}{dx^2} - \frac{dz}{dx} - 2y = 4x,$$

$$\frac{dy}{dx} + 4\frac{dz}{dx} - 3z = 0.$$

7. If a shaft has masses attached at its two ends and is given an initial twist then left free to vibrate under the action of its loads and strains, the differential equations of motion may be written

$$J_o\frac{d\theta_o}{dt} + J_1\frac{d\theta_1}{dt} = k,$$

$$J_o\frac{d^2\theta_o}{dt^2} = \frac{IG}{L}(\theta_1 - \theta_o),$$

where J_o and J_1 are the moments of inertia of the attached masses about the shaft axis, θ_o and θ_1 are the angular displacements of these masses, I is the moment of inertia of the section area of the shaft, G is the modulus of rigidity of the shaft, L the length of shaft, and k is a constant depending on the initial angular velocity. Find θ_1 and θ_o as functions of t.

8. If a disturbing moment, M sin αt, is applied to the mass at one end of the shaft in example 7, the differential equations may be written

$$J_o\frac{d^2\theta_o}{dt^2} = \frac{IG}{L}(\theta_1 - \theta_o) + M\sin\alpha t,$$

$$J_1\frac{d^2\theta_1}{dt^2} = -\frac{IG}{L}(\theta_1 - \theta_o).$$

Solve these equations, discriminating between the case when

$$\alpha^2 = \frac{IG\,(J_o + J_1)}{L\,J_o\,J_1}$$

and the case when this relation is not fulfilled.

9. The differential equations for the currents, i_1 and i_2, in two inductively coupled electric circuits are

$$R_1 i_1 + L_1 \frac{di_1}{dt} + \frac{1}{K_1}\int i_1 dt - M\frac{di_2}{dt} = E\cos\omega t,$$

$$R_2 i_2 + L_2 \frac{di_2}{dt} + \frac{1}{K_2}\int i_2 dt - M\frac{di_1}{dt} = 0,$$

where R_1, L_1 and K_1 are resistance, inductance and capacity, respectively of the first circuit, R_2, L_2 and K_2 are the corresponding constants for the second circuit, M is mutual inductance between the circuits and E cos ωt is a periodic electromotive force impressed upon the first circuit. If $E = 0$ and the resistances may be neglected, show that the current in each circuit is given by an equation of the form

$$i = C_1 \sin(\omega_1 t + \phi_1) + C_2 \sin(\omega_2 t + \phi_2)$$

where C_1, C_2, ϕ_1 and ϕ_2 are arbitrary constants and ω_1 and ω_2 depend on the constants of the circuits. (Let $q_1 = \int i_1 dt$ and $q_2 = \int i_2 dt$, and assume $L_1 L_2 > M^2$).

10. If E is not zero in 9, what is the character of terms to be added to the expressions for current in the two circuits. Discuss particularly the condition when ω is equal to ω_1 or ω_2, neglecting resistances.

11. Without neglecting resistances, write the fourth order differential equations for q_1 and q_2 that arise in solving the system in 9.

12. Suppose a body suspended by a coiled spring and a second body suspended from the first by a second coiled spring. Suppose the bodies to oscillate in a vertical line, not swinging. Write differential equations for the motions of the two bodies, assigning all necessary notations. Neglect mass and weight of springs, and resistance of air. Solve the differential equations.

13. Solve the system

$$\frac{dx}{dt} + \frac{dy}{dt} + 5y = t,$$

$$\frac{d^2x}{dt^2} - y = 0,$$

$$\frac{dy}{dt} - \frac{dz}{dt} + y = e^t.$$

14. Solve the system

$$\frac{dx}{y^2t} = \frac{dy}{x^2t} = \frac{dt}{xy^2}.$$

15. Solve the system

$$\frac{dx}{-y} = \frac{dy}{x} = \frac{dt}{1 + t^2}.$$

16. Solve the system

$$\frac{dx}{dt} = \frac{y + t}{x + y}, \quad \frac{dy}{dt} = \frac{x + t}{x + y}.$$

17. Solve the system

$$\frac{dx}{z} = \frac{dy}{zy} = \frac{dz}{x} = \frac{dt}{tz}.$$

18. Solve by reducing to a system of the first order the equation

$$\frac{d^3y}{dx^3} + \left(\frac{d^2y}{dx^2}\right)^2 = 1.$$

Art. 70. Total Differential Equations

1. **Condition of Integrability.** An equation of the form

$$Pdx + Qdy + Rdz = 0 \tag{1}$$

where P, Q and R are functions of x, y and z, is called a **total differential equation.** The form may be extended to any number of variables, but we shall deal at present with the case of three variables.

Consider now a relation of the form

$$f(x, y, z) = C \tag{2}$$

with the consequent equation

$$\frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz = 0. \tag{3}$$

If (2) is the solution of (1) then there must exist a function $\phi(x, y, z)$ such that when (1) is multiplied by this function it reduces to (3). In this case we should have

$$\frac{\partial f}{\partial x} = \phi P, \quad \frac{\partial f}{\partial y} = \phi Q, \quad \frac{\partial f}{\partial z} = \phi R. \tag{4}$$

Now assuming that

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x'}, \quad \frac{\partial^2 f}{\partial y\,\partial z} = \frac{\partial^2 f}{\partial z\,\partial y'}, \quad \frac{\partial^2 f}{\partial z\,\partial x} = \frac{\partial^2 f}{\partial x\,\partial z'}, \tag{5}$$

we must have

$$\begin{aligned}
\phi\frac{\partial P}{\partial y} + P\frac{\partial \phi}{\partial y} &= \phi\frac{\partial Q}{\partial x} + Q\frac{\partial \phi}{\partial x'}\\
\phi\frac{\partial Q}{\partial z} + Q\frac{\partial \phi}{\partial z} &= \phi\frac{\partial R}{\partial y} + R\frac{\partial \phi}{\partial z'}\\
\phi\frac{\partial R}{\partial x} + R\frac{\partial \phi}{\partial z} &= \phi\frac{\partial P}{\partial z} + P\frac{\partial \phi}{\partial z}.
\end{aligned} \tag{6}$$

Here we have three equations to be fulfilled by a single unknown function ϕ, and this is possible only in case the functions P, Q and R are specially related. If we multiply the three equations by R, P and Q respectively, and add, then divide by ϕ we find

$$P\left(\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y}\right) + Q\left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) + R\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = 0. \tag{7}$$

If this relation is not satisfied no multiplier ϕ can exist to reduce (1) to (3), hence no solution of form (2) can exist. Equation (7) is the **condition of integrability** for the case of three variables.

If condition (25) of Art. 57 is fulfilled then (1) is already in form (3), and the function ϕ may be taken as unity. It is not convenient, in general, to solve by finding a function ϕ to serve as integrating factor.

For the case of four variables we may write in place of (1) the equation

$$Pdx + Qdy + Rdz + Sdt = 0, \tag{8}$$

and there will be four equations of form (4), and therefore six equations of form (5), one for each pair of variables selected from the four variables. One equation of form (6) comes from each equation of form (5), each associated with a certain pair of variables. Each variable appears in three different pairs, and if we omit these pairs and consider equations (6) that correspond to the remaining three pairs we find an equation of form (7). Doing this for the four variables in turn we find four equations of form (7), but only three of these are independent. The fourth equation of form (7) will be satisfied as a result of the first three,

as will any other relation among P, Q, R and S and their partial derivatives that may be deduced from the equations of form (6). Any three of the four equations of form (7) constitute the **conditions of integrability** for the case of four variables. The first of the four equations may be written exactly as (7), and the other three by putting S and t in place of R and z, then in place of Q and y, then in place of P and x.

2. **Solution of equations that are integrable.** If an equation of form (1) satisfies the condition (7) we may proceed as follows to find the solution in form (2). Write the equation

$$Pdx + Qdy = 0 \tag{9}$$

and solve as a differential equation in x and y, regarding z as a parameter. The arbitrary constant that appears in the solution may be taken as a function of the parameter z, thus the solution may be written

$$F(x, y, z) = f(z). \tag{10}$$

Now if we differentiate this with respect to all three variables we have

$$\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \left(\frac{\partial F}{\partial z} - \frac{df}{dz}\right)dz = 0. \tag{11}$$

There must exist a function $\phi(x, y, z)$ such that

$$\frac{\partial F}{\partial x} = \phi P, \quad \frac{\partial F}{\partial y} = \phi Q,$$

so that, (11) will be (1) multiplied by ϕ if

$$\frac{\partial F}{\partial z} - \frac{df}{dz} = \phi R. \tag{12}$$

It may be shown that if the condition of integrability is satisfied then equation (12) gives a differential equation in f and z only, and f may be found as a function of z, involving an arbitrary constant. With f(z) so determined, (10) is a solution of (1).

As an example consider

$$yzdx - zxdy - y^2dz = 0.$$

The condition of integrability gives

$$yz(-x + 2y) - zx(-y) - y^2(z + z) = 0,$$

and this being satisfied we consider the equation

$$yzdx - zxdy = 0,$$

regarding z as constant. The solution may be written

$$\frac{x}{y} = f(z),$$

from which

$$\frac{dx}{y} - \frac{xdy}{y^2} - \frac{df}{dz}\,dz = 0.$$

Comparing this with the original equation we find

$$\frac{df}{dz} = \frac{1}{z},$$

therefore, $f(z) = \log z + C$, and the solution may be written in either of the forms

$$\frac{x}{y} = \log z + C, \quad z = Ce^{\frac{x}{y}}.$$

To solve equation (8) we should write (9) as above, and take the arbitrary constant as a function of z and t, so that instead of (10) we should have

$$F(x, y, z, t) = f(z, t).$$

Differentiating this with respect to all four variables and comparing with (8), we should find an equation involving the three variables, z, t and f, providing that the three conditions of integrability are satisfied by the given equation. Also this new equation in three variables will be integrable and may be solved as equation (1).

In case of an integrable equation in three variables in which P, Q and R are **homogeneous and of the same degree,** we may make the substitution $x = uz$, $y = vz$ and obtain a new equation in u, v and z, and if the coefficient of dz is not zero we may write

$$P_1du + Q_1dv + \frac{dz}{z} = 0, \tag{13}$$

where $P_1du + Q_1dv$ is an exact differential, say the differential of a function f(u, v), and the solution is

$$f(u, v) + \log z = C. \tag{14}$$

For the similar case in four variables we may put $x = ut$, $y = vt$, $z = wt$, and obtain a new equation

$$P_1du + Q_1dv + R_1dw + \frac{dt}{t} = 0, \tag{15}$$

in which the first three terms give an exact differential and the solution may be written

$$f(u, v, w) + \log t = C. \tag{16}$$

As an illustration take

$$(y^2 + z^2)\,dx + xzdy - xydz = 0.$$

The condition of integrability is satisfied, and P, Q and R are homogeneous, and all of the second degree. Putting $x = uz$, $y = vz$ we have

$$z^2(v^2 + 1)\,(udz + zdu) + z^2u(vdz + zdu) - z^2vudz = 0,$$

and after dividing by z^2 and collecting terms,

$$z(v^2 + 1)\,du + zudv + (uv^2 + u)\,dz = 0,$$

which reduced to

$$\frac{du}{u} + \frac{dv}{v^2 + 1} + \frac{dz}{z} = 0.$$

This is in form (13), and the solution in form (14) is

$$\log u + \tan^{-1} v + \log z = C.$$

Replacing u and v by their values in x, y and z we may write the solution in the form,

$$\log x + \tan^{-1} \frac{y}{z} = C.$$

In arriving at (13) we suppose the coefficient of dz to be different from zero. If this coefficient should be zero, the form (13) cannot be reached, but the equation becomes a differential equation in u and v only, and may be solved as in Chapter XIV.

For example, consider

$$(x + z)dx + (2z - x)dy - (x + 2y)dz = 0.$$

The condition of integrability is satisfied and the coefficients of the differentials are homogeneous and all of the first degree. Putting $x = uz$ and $y = vz$ we have

$$z(v + 1)\,(udz + zdu) + z(2 - u)\,(vdz + zdv) - z(u + 2v)\,dz = 0,$$

and after dividing by z and collecting coefficients,

$$z(v + 1)du + z(2 - u)dv = 0,$$

the coefficient of dz reducing to zero for this special numerical case. The solution is

$$\frac{2 - u}{v + 1} = C = \frac{2z - x}{y + z}.$$

3. **Equations not integrable.** As we have seen above, if a total differential equation is integrable its general solution consists of a single equation among the variables, involving an arbitrary constant. In the case of three variables this solution represents a one-parameter family of surfaces such that the normal line at any point (x_o, y_o, z_o) of any of these surfaces has direction cosines proportional to the three values $P(x_o, y_o, z_o)$, $Q(x_o, y_o, z_o)$, $R(x_o, y_o, z_o)$. If the given equation in three variables does not satisfy the condition of integrability, then no such family of surfaces exists. If we assume any relation among the variables, as

$$F(x, y, z) = 0, \tag{17}$$

we may express one variable in terms of the other two, and thus reduce (1) to an equation in two variables, say x and y, and if the solution of this is

$$f(x, y, C) = 0 \tag{18}$$

then (17) and (18) together give a solution of equation (1). Geometrically, (17) and (18) determine a family of space curves, and if (1) is integrable these curves lie on the surfaces given by the solution in form (2). Of course (17) and (2) determine the curves, but (17) being a purely arbitrary relation not depending on (1), the curves may be regarded as arbitrarily drawn on the surfaces (2), and contribute nothing to the information contained in (2). On the other hand, (17) and (18) of themselves can never lead to a solution of form (2). In case (1) is not integrable, however, (17) and (18) afford the only form of solution available.

In the case of an equation in four variables that does not satisfy the conditions of integrability we may assume two relations

$$F_1(x, y, z, t) = 0, \quad F_2(x, y, z, t) = 0, \tag{19}$$

and express two of the variables in terms of the other two, thus reducing (8) to an equation in two variables, say x and y, and writing the solution of this in form (18), we have in (18) and (19) a solution of the given equation.

If by inspection a single relation among the four variables may be assumed which reduces (8) to an integrable equation in three variables then the solution of this together with the assumed relation will give a solution depending on one assumed relation, more desirable than a solution depending on two assumed relations.

For example, take

$$tdx - xdy + ydz - zdt = 0.$$

The conditions of integrability are not satisfied, but if we assume

$$t - y = 0, \quad z - x = 0,$$

the equation reduces to

$$2ydx - 2xdy = 0,$$

and the solution of this is

$$x - Cy = 0.$$

The relations

$$t - y = 0, \quad z - x = 0, \quad x - Cy = 0,$$

satisfy the given equation.

4. **Systems of total differential equations.** Let us consider two simultaneous equations in three variables

$$\begin{aligned} P_1dx + Q_1dy + R_1dz &= 0, \\ P_2dx + Q_2dy + R_2dz &= 0. \end{aligned} \tag{20}$$

If these equations are each integrable, then the two solutions, each involving an arbitrary constant, constitute the solution of the system (20). Geometrically the solution gives a two-parameter family of space curves, the intersections of two families of surfaces.

If only one of the equations is integrable, the solution of this one gives a relation among the variables by which the other equation may be reduced to an equation in two variables. The solution of this gives a second relation which, together with the first, furnishes the solution of the system.

Algebraically the system (20) is equivalent to the system

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R'} \tag{21}$$

where $P = Q_1R_2 - Q_2R_1$, $Q = R_1P_2 - P_1R_2$ and $R = P_1Q_2 - P_2Q_1$. A solution of the system (21) by the method of Art. 69 will be a solution of the system (20). In case both equations of (20) fail to satisfy the condition of integrability the use of (21) is the only means of finding a solution.

A system of three equations of form (8) yields a solution consisting of three relations, each involving an arbitrary constant, providing that at least two of the equations are integrable. In any case the system leads algebraically to a system of type (20) of Art. 69 which may be solved by the methods of that article.

A system of two equations of form (8) yields a solution consisting of two relations, each involving an arbitrary constant, providing that both equations are integrable. Otherwise a relation may be assumed which reduces the equations to three variables and thus to a system of form (21). The solution of this system together with the assumed relation constitutes a solution of the given system.

In the preceding paragraphs we see that a system of total differential equations is equivalent to a system of differential equations of the first order, and advantage may be taken of this fact in finding the solution of the system. This advantage is reciprocal, for equation (21) of Art. 69 is, in fact, a total differential equation. If by any choice of the multipliers X, Y, Z and T the equation may be made integrable, then the solution of this total differential equation furnishes one of the relations sought in the solution of the given system.

It should be noted that in dealing with a system of total differential equations any equation of the system may be replaced by an algebraic combination of this equation with one or more other equations of the system, as the sum or difference of multiples of the equations. New equations formed this way may be integrable when the original equations are not integrable.

EXERCISES LXVII

1. Show that the following equations satisfy the condition of integrability.

$$xydx + x^2dy + xzdz = 0,$$

$$xdx + zdy + ydz = 0.$$

2. Show that the sum or the difference of the two equations written above is an equation which is not integrable.

3. Show that solutions for the equations written in 1 are respectively $x^2 + 2yz = C$ and $2xy + z^2 = C$.

4. Show that the following equation is integrable,

$$t^3ydx - xt^3dy + ty^2dz - 2y^2zdt = 0.$$

Also show that the solution is

$$\frac{x}{y} + \frac{z}{t^2} = C.$$

5. Solve the equation

$$(y + z)dx + dy + dz = 0.$$

6. Solve the equation

$$(y^2 + yz)dx + (yz + z^2)dy + (y^2 - xy)dz = 0.$$

7. Show that the following equation does not satisfy the condition of integrability, and find its solution assuming that $x + y + z = A$.

$$(x + y)dx + (x - y)dy + (x + y)dz = 0.$$

8. Solve the equation

$$(z + 1)\,(xdy + ydx) - (x^2 + y^2)dz = 0.$$

9. Show that the equation

$$C_v vdp + C_v kpdv - Bdq = 0$$

does not satisfy the condition of integrability if $k > 1$. By comparing this equation with equation (28) of Art. 57 show that the quantity of heat cannot be expressed as a function of p and v.

10. Show that the solution of the equation in 9 under the assumption that $pv = C$ is

$$q = \frac{C_v}{B}(k - 1)C \log_e v + c,$$

and thus establish equation (33) of Art. 57.

11. Solve the equation in 9, under the assumption that $q = a$, and thus establish equation (34) of Art. 57 as the condition for adiabatic expansion.

12. By associating the total differential equation

$$C_v \frac{dp}{p} + C_v k \frac{dv}{v} - dS = 0$$

with equation (31) of Art. 57, establish equation (32) of that article as the equation for entropy.

13. Show that if the equation

$$Xdx + Ydy + Zdz + dU = 0$$

is integrable and if X, Y and Z are functions of x, y and z only, the sum of the first three terms is an exact differential.

14. Find U as a function of x, y and z from the equation in 13 if $X = yz$, $Y = zx + y$ and $Z = xy + z^2$.

15. Let the equation in 13 be as specified in 14 and find the surfaces on which U is constant (equipotential surfaces).

16. If in 13, $X = x$, $Y = y$, $Z = x + y$, show that the equation is not integrable. Solve the equation under the assumptions $x = z$, $y = z^2$, and thus find the work done against the force field in going from the origin to the point (1, 4, 2) along the path given by the assumed equations.

17. Solve the system of equations

$$(x + y)dx + (x - y)dy + (x + y)dz = 0,$$
$$(x - y)dx + (x + y)dy + (x - y)dz = 0.$$

18. Solve the system of equations

$$(x + xy)dx + (x^2 + z)dy + (y + xz)dz = 0,$$
$$(xy - x)dx + (x^2 - z)dy + (xz - y)dz = 0.$$

(Compare with sum and difference of equations in 1.)

19. Solve the system

$$\frac{dx}{3y - 2z} = \frac{dy}{z - 3x} = \frac{dz}{2x - y}.$$

20. Solve the system

$$tdx + zdy + tdz + ydt = 0,$$
$$tdy + zdz - t^2dt = 0,$$
$$dz - 2tdt = 0.$$

BOOKS FOR SUPPLEMENTARY READING

Advanced Calculus, W. F. Osgood (Macmillan, 1925), Chapters XIV and XV.
Advanced Calculus, E. B. Wilson (Ginn, 1912), Chapters VII–X.
Differential Equations, H. B. Phillips (Wiley, 1922).
Differential Equations in Applied Chemistry, Hitchcock and Robinson (Wiley, 1923).
Differential Equations, A. Cohen (Heath, 1906), Chapters I–X.
Differential Equations, D. A. Murray (Longmans, 1919), Chapters I–XI.
Differential Equations, H. Bateman (Longmans, 1918), Chapters I–V and VII.
Differential Equations, A. R. Forsyth (Macmillan, 1903), Chapters I–IV and VIII.

CHAPTER XVII

PARTIAL DIFFERENTIAL EQUATIONS

Art. 71. Meaning of Partial Differential Equations

In the preceding article we saw that a single equation in three or more variables involving one arbitrary constant gives rise to a total differential equation. Consider now an equation in three variables involving two or more arbitrary constants,

$$f(x, y, z, C_1, C_2, \ldots, C_n) = 0. \tag{1}$$

We may regard this as an equation defining z as a function of two independent variables, x and y, and we may take partial derivatives with respect to x and y,

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0, \tag{2}$$

$$\frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0, \tag{3}$$

and, if necessary, these may be differentiated, giving three new equations, involving the partial derivatives of z of the second order. Further differentiation will bring in higher derivatives, and we may proceed until we have in all n + 1 equations, so that the n arbitrary constants may be eliminated, the result being an equation involving x, y and z and certain partial derivatives of z with respect to x and y. Such an equation is called a **partial differential equation,** and the equation (1) from which it is derived is its primitive.

For example, let (1) take the form

$$z - (x + a)(y + b) = 0, \tag{4}$$

where a and b are arbitrary constants. Now (2) and (3) become

$$-(y + b) + \frac{\partial z}{\partial x} = 0, \tag{5}$$

$$-(x + a) + \frac{\partial z}{\partial y} = 0, \tag{6}$$

and from the three equations we at once obtain the differential equation

$$z - \frac{\partial z}{\partial x}\frac{\partial z}{\partial y} = 0. \tag{7}$$

It will be convenient in what follows to assign the notations

$$p = \frac{\partial z}{\partial x},\ q = \frac{\partial z}{\partial y},\ r = \frac{\partial^2 z}{\partial x^2},\ s = \frac{\partial^2 z}{\partial x\,\partial y},\ t = \frac{\partial^2 z}{\partial y^2}.$$

Suppose (1) to take the form

$$ax + by + cz + gx^2 + hy^2 + kz^2 = 0. \tag{8}$$

Taking derivatives we have

$$a + cp + 2gx + 2kzp = 0,$$
$$b + cq + 2hy + 2kzq = 0,$$

and taking derivatives again,

$$cr + 2g + 2kzr + 2kp^2 = 0,$$
$$cs + 2kzs + 2kpq = 0,$$
$$ct + 2h + 2kzt + 2kq^2 = 0.$$

We have, in all, six equations, homogeneous in the constants, and eliminating the constants we have

$$\begin{vmatrix} x & y & z & x^2 & y^2 & z^2 \\ 1 & 0 & p & 2x & 0 & 2zp \\ 0 & 1 & q & 0 & 2y & 2zq \\ 0 & 0 & r & 2 & 0 & (2zr + 2p^2) \\ 0 & 0 & s & 0 & 0 & (2zs + 2pq) \\ 0 & 0 & t & 0 & 2 & (2zt + 2q^2) \end{vmatrix} = 0 \tag{9}$$

The expansion of this determinant gives the desired differential equation.

Again, let u and v be known functions of x, y and z, and consider an arbitrary relation between them,

$$\phi(u, v) = 0. \tag{10}$$

This equation may be regarded as defining z as a function of x and y, and, taking partial derivatives, we have

$$\frac{\partial \phi}{\partial u}\left(\frac{\partial u}{\partial x} + p\frac{\partial u}{\partial z}\right) + \frac{\partial \phi}{\partial v}\left(\frac{\partial v}{\partial x} + p\frac{\partial v}{\partial z}\right) = 0,$$

$$\frac{\partial \phi}{\partial u}\left(\frac{\partial u}{\partial y} + q\frac{\partial u}{\partial z}\right) + \frac{\partial \phi}{\partial v}\left(\frac{\partial v}{\partial y} + q\frac{\partial v}{\partial z}\right) = 0. \tag{11}$$

The equations permit us to eliminate $\frac{\partial \phi}{\partial u}$ and $\frac{\partial \phi}{\partial v}$ and obtain an equation of the form

$$Pp + Qq = R \tag{12}$$

where P, Q and R are known functions of x, y and z. The arbitrary relation (10) may be written in various forms, and if more than one arbitrary function is involved it may be necessary to extend the system (11) by further differentiation before the arbitrary functions may be eliminated, but in any case, a single relation involving one or more arbitrary functions gives rise to a partial differential equation of the first or higher order.

To illustrate, let $u = x + z$ and $v = yz$, then equations (11) become

$$\frac{\partial \phi}{\partial u}(1 + q) + \frac{\partial \phi}{\partial v}(0 + py) = 0,$$

$$\frac{\partial \phi}{\partial u}(0 + q) + \frac{\partial \phi}{\partial v}(z + qy) = 0,$$

and we have at once

$$(1 + p)(z + qy) - pqy = 0$$

or

$$zp + yq = -z.$$

To illustrate further, consider the arbitrary relation $z = \phi(x^2 + y^2)$. Taking partial derivatives,

$$p = 2x\phi', \quad q = 2y\phi',$$

therefore we have the partial differential equation

$$yp - xq = 0.$$

As a final illustration take

$$z = \phi_1(x + y) + \phi_2(x - y),$$

taking first and second partial derivatives we obtain

$$p = \phi_1' + \phi_2',\ q = \phi_1' - \phi_2',\ r = \phi_1'' + \phi_2'',\ t = \phi_1'' + \phi_1''.$$

The desired partial differential equation is of the second order, $r = t$.

We see that partial differential equations may arise from a relation (1) involving two or more arbitrary constants, also that they may arise from a relation involving one or more arbitrary functions. Moreover, the same equation may rise from the two different primitives. For example, the equation $r = t$ derived in the final illustration above may be derived from the primitive $z = ax^2 + ay^2 + bx + cy$. Now this is a very special case of the form used above, obtained by writing

$$\phi_1(x + y) = A(x + y)^2 + B(x + y),$$
$$\phi_2(x - y) = A(x - y)^2 + C(x - y),$$

where $2A = a$, $B + C = b$, $B - C = c$.

A **solution** of a partial differential equation is any relation among the variables that is consistent with the given equation. A primitive, involving arbitrary constants, from which the given equation may be derived is a **complete** solution, while the primitive, involving arbitrary functions, from which the given equation may be derived, is a **general solution.** An equation may have an indefinite number of complete solutions, all distinct, and all special cases under the general solution.

As in previous chapters we have used the arbitrary constants to satisfy initial conditions that may be specified along with the ordinary differential equation, so here we may use the arbitrary features of the solution to satisfy certain conditions that may be known along with the partial differential equation. These conditions may pertain to the type of the function, as periodic, exponential, etc., or they may specify **boundary values,** that is, the surface determined by the solution may be required to pass through a certain fixed curve, or satisfy other specified conditions.

For example, suppose it required to solve the equation $r = t$ subject to the conditions that z is a harmonic function of x when y is constant, and that $z = 0$ for all values of y when $x = 0$. The harmonic character is obtained by assigning the two arbitrary functions written above the values

$$\phi_1(x + y) = a_1 \sin \omega(x + y) + b_1 \cos \omega(x + y),$$
$$\phi_2(x - y) = a_2 \sin \omega(x - y) + b_2 \cos \omega(x - y).$$

When $x = 0$ the solution reduces to

$$z = (a_1 - a_2) \sin \omega y + (b_1 + b_2) \cos \omega y,$$

and this will be zero for all values of y if $a_1 = a_2$ and $b_1 = -b_2$, so that the solution which satisfies all conditions may be written

$$\begin{aligned} z &= a [\sin \omega(x + y) + \sin \omega(x - y)] \\ &\quad + b [\cos \omega(x + y) - \cos \omega(x - y)] \\ &= 2a \sin \omega x \cos \omega y - 2b \sin \omega x \sin \omega y \\ &= 2 \sin \omega x (a \cos \omega y - b \sin \omega y). \end{aligned}$$

The three constants, a, b and ω remain arbitrary and may be assigned values to satisfy further restrictions.

Again, suppose it required to solve the same equation subject to the conditions that z shall be a linear function of x and y and that $z = 1 - y$ when $x = 1$. Here we may take

$$\begin{aligned} z &= a(x + y) + b(x - y) + c \\ &= (a + b)x + (a - b)y + c. \end{aligned}$$

When $x = 1$ this reduced to

$$z = a + b + c + (a - b)y,$$

and we must have $a + b + c = 1$, $a - b = -1$, so that $b = 1 + a$, $c = -2a$, and the required solution is

$$z = (2a + 1)x - y - 2a.$$

The constant a remains arbitrary.

EXERCISES LXVIII

1. Find the partial differential equations of which the following are primitives, considering a, b and c as arbitrary.

(a) $z = a(x - y) + b$;

(b) $z = ax + by + ab$;

(c) $(x - a)^2 + (y - b)^2 + z^2 = c^2$.

2. Find the partial differential equations of which the following are primitives, f and ϕ being symbols of arbitrary functions.

(a) $\phi(xy - z, x^2 + y^2) = 0$; (d) $z = f(x + 2y) + \phi(x - 2y)$;

(b) $\phi(x + y - z, x - y) = 0$; (e) $z = x + \phi(x^2 + y^2 + z^2)$;

(c) $z = e^{3y}\phi(x - y)$; (f) $z = \phi(x + y) + f(xy)$.

3. Determine the arbitrary function in 2(b) so that the point (x, y, z) lies on a plane whose trace on the plane $x = 0$ is $z = 2y - 5$.

4. Determine the arbitrary function in 2(e) so that the point (x, y, z) shall lie on a sphere, and so that the trace of this sphere on the plane $y = 0$ is a circle of unit radius.

5. Determine the arbitrary functions in 2(f) so that z contains only even powers of x.

Art. 72. Partial Differential Equations of First Order

1. Linear equations of the first order. An equation of the first degree in the partial derivatives p and q,

$$Pp + Qq = R, \tag{1}$$

where P, Q, and R are functions of x, y and z, is called a **linear** equation. It may be shown that if u and v are functions of x, y and z such that $u = C_1$ and $v = C_2$ constitute the solution of the system

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R} \tag{2}$$

then we may write the general solution of (1) as

$$\phi(u, v) = 0, \tag{3}$$

where ϕ is an arbitrary function.

Any term in (2) is equal to the ratio

$$\frac{\dfrac{\partial u}{\partial x}dx + \dfrac{\partial u}{\partial y}dy + \dfrac{\partial u}{\partial z}dz}{P\dfrac{\partial u}{\partial x} + Q\dfrac{\partial u}{\partial y} + R\dfrac{\partial u}{\partial z}}, \tag{4}$$

but if $u = C_1$ the numerator is of (4) zero, and hence also the denominator. Further we have

$$\frac{\partial u}{\partial x} + p\frac{\partial u}{\partial z} = 0, \quad \frac{\partial u}{\partial y} + q\frac{\partial u}{\partial z} = 0,$$

therefore the denominator equated to zero is equivalent to (1). Similarly $v = C_2$ is shown to satisfy (1).

Now if $u = C_1$ and $v = C_2$ satisfy (2) then $\phi(u, v)$ when put for u in the denominator of (4) will cause it to be zero, hence furnishes a solution of (1). Thinking of ϕ as a function of x, y and z we have

$$\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \phi}{\partial v}\frac{\partial v}{\partial x},$$

$$\frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial \phi}{\partial v}\frac{\partial v}{\partial y},$$

$$\frac{\partial \phi}{\partial z} = \frac{\partial \phi}{\partial u}\frac{\partial u}{\partial z} + \frac{\partial \phi}{\partial v}\frac{\partial v}{\partial z}.$$

Multiplying these respectively by P, Q and R and adding we see that if the denominator of (4) vanishes for u and v then it does for ϕ.

For example take

$$yp - xq = 1.$$

We solve the system

$$\frac{dx}{y} = -\frac{dy}{x} = dz.$$

From the first two we have $x^2 + y^2 = C_1$, and from

$$\frac{ydx - xdy}{x^2 + y^2} = dz$$

we have $\tan^{-1}\frac{y}{x} + z = C_2$, therefore the general solution of the given equation is

$$\phi\left(x^2 + y^2, \tan^{-1}\frac{y}{x} + z\right) = 0.$$

2. **First order equations that are not linear.** For equations that are not linear it is not generally possible to write the general solution. If, however, we are able to find a complete solution we may employ it to obtain any number of distinct solutions that are also complete. Let

$$F(x, y, z, p, q) = 0 \tag{5}$$

be the partial differential equation, and let a complete solution be written

$$f(x, y, z, a, b) = 0. \tag{6}$$

If we form the two equations

$$\frac{\partial f}{\partial a} = 0, \qquad \frac{\partial f}{\partial b} = 0, \tag{7}$$

and eliminate a and b from (6) and (7) we shall have a solution of (5). To see this we may regard (6) and (7) as determining z, a and b as functions of x and y, then differentiating (6) with respect to x and y we have

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} p + \frac{\partial f}{\partial a}\frac{\partial a}{\partial x} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial x} = 0,$$

$$\frac{\partial f}{\partial y} + \frac{\partial f}{\partial z} q + \frac{\partial f}{\partial a}\frac{\partial a}{\partial y} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial y} = 0,$$

and taking account of (7) these reduce to the same forms obtained when regarding a and b as constants, and, together with (6) determine p, q and z so as to satisfy (5).

The solution found in this way contains no arbitrary constant and is called a **singular solution** of (5).

Now let us assume a relation between the two constants in (6), say b is a function of a, then we may replace (7) by

$$b = \phi(a), \quad \frac{\partial f}{\partial a} + \phi'(a)\frac{\partial f}{\partial b} = 0. \tag{8}$$

Proceeding as above, it may be shown that elimination of a and b from (6) and (8) gives a solution of (5). Moreover, since $\phi(a)$ is an arbitrary function, the solution may involve several constants. We may regard (6) and (8) together as giving the general solution, but since the elimination of a and b involves a choice of the function ϕ, we are able to write down only particular solutions.

As illustration take the partial differential equation $z = px + qy - pq$. A complete solution is $z - ax - by + ab = 0$, and equations (7) give $-x + b = 0$, $-y + a = 0$, and eliminating a and b we have the singular solution $z = xy$. If we take $\phi(a) = ma + C$, then equations (8) give

$$b = ma + C, \; -x + b + m(-y + a) = 0,$$

and elimination of a and b gives the solution

$$4mz - (x + my)^2 + 2C(x - my) - C^2 = 0.$$

Again, if we take $\phi(a) = \frac{C^2}{a}$ we have for (8) the equations

$$ab = C^2, \quad -x + b - \frac{C^2}{a^2}(-y + a) = 0,$$

and elimination of a and b yields the solution

$$z - 2Cx^{\frac{1}{2}}y^{\frac{1}{2}} + C^2 = 0.$$

In the following paragraphs we present briefly certain methods for writing complete solutions for partial differential equations of the types specified. The method of the final paragraph is of general application, but often more difficult to apply than the special methods which precede it.

(a) **The form** $f(p, q) = 0$. Here x, y and z do not appear explicitly in the differential equation. A complete solution is given by

$$z = ax + by + C, \quad f(a, b) = 0.$$

For example, take $p^2 + q + 1 = 0$, then we must have $a^2 + b + 1 = 0$ and the complete solution is $z = ax - (1 + a^2)y + C$.

(b) **The form** $z = px + qy + f(p\,q)$. This is similar to the ordinary differential equation known as Clairaut's equation. The complete solution is

$$z = ax + by + f(a, b).$$

(c) **The form** $f(z, p, q) = 0$. Here x and y do not occur explicitly. We may try $z = \phi(u)$, where $u = x + ay$. Now we shall have

$$p = \phi'(u)\frac{\partial u}{\partial x} = \frac{dz}{du}, \quad q = \phi'(u)\frac{\partial u}{\partial y} = a\frac{dz}{du}.$$

Substituting these values in the given equation we obtain an ordinary differential equation of the first order in u and z. Solving this we have z as a function of u and a and a new constant, and when u is replaced by $x + ay$ we have a complete solution.

Take, for example, $pz + q^2 = 0$. The ordinary differential equation is

$$z\frac{dz}{du} + a^2\left(\frac{dz}{du}\right)^2 = 0,$$

giving the solution

$$z = Ce^{-\frac{u}{a^2}},$$

and the complete solution of the given equation may be written

$$z = Ce^{-\frac{x+ay}{a^2}}.$$

(d) **The forms** $f_1(x, y, z, p) = 0$, $f_2(x, y, z, q) = 0$. In these forms only one of the partial derivatives occurs. To solve the first we may regard y as constant and treat the equation as an ordinary differential equation in x and z, but take the arbitrary constant as an arbitrary function of y. This method yields the general solution, not merely a complete solution. The second form is treated as the first, but x and y change places in the discussion.

For example, consider $x^2p^2 - (y - z)^2 = 0$. We may write

$$\frac{dz}{y - z} \pm \frac{dx}{x} = 0,$$

and the solution takes one of the forms

$$\frac{x}{y - z} = \phi(y), \qquad x(y - z) = \phi(y).$$

(e) **The form** $f_1(x, p) = f_2(y, q)$. Here z does not occur explicitly, and the terms in x and p are separated from those in y and q. We may write

$$f_1(x, p) = a, \quad f_2(y, q) = a$$

and solving these algebraically for p and q then integrating we obtain two expressions for z,

$$z = F_1(x, a) + \phi_1(y),$$

$$z = F_2(y, a) + \phi_2(x),$$

where F_1 and F_2 are known functions and ϕ_1 and ϕ_2 are arbitrary. The two forms agree if $\phi_1 = F_2 + b$ and $\phi_2 = F_1 + b$, so that

$$z = F_1(x, a) + F_2(y, a) + b.$$

This is a complete solution of the given equation.

(f) **The general form** $f(x, y, z, p, q) = 0$. We are to find z as a function of x and y, and evidently we shall have

$$dz = pdx + qdy$$

and if p and q were known as functions of x, y and z we should be able to treat this as a total differential equation. We shall be able to find p and q if we can find a second relation

$$F(x, y, z, p, q) = 0$$

to combine with the given equation. It may be shown that any relation, involving at least one of the partial derivatives p and q, found from the system

$$\frac{dp}{\frac{\partial f}{\partial x} + p\frac{\partial f}{\partial z}} = \frac{dq}{\frac{\partial f}{\partial y} + q\frac{\partial f}{\partial z}} = -\frac{dx}{\frac{\partial f}{\partial p}} = -\frac{dy}{\frac{\partial f}{\partial q}}$$

will serve as the second relation. The simpler this second relation the better, but it should involve one arbitrary constant. The second arbitrary constant arises in solving the total differential equation. We omit here the derivation of the method, but illustrate by an example the use of the method.

Take the equation $(p^2 + q^2)y - qz = 0$. The system indicated above becomes

$$\frac{dp}{-pq} = \frac{dq}{p^2} = -\frac{dx}{2py} = -\frac{dy}{2qy - z}.$$

The first two members give the relation $p^2 + q^2 = a^2$, where a^2 is the arbitrary constant. Using this with the given equation to solve for p and q we have

$$p = \frac{a}{z}\sqrt{z^2 - a^2y^2}, \qquad q = a^2\frac{y}{z}.$$

We have now to solve the total differential equation

$$dz = \frac{a}{z}\sqrt{z^2 - a^2y^2}\,dx + a^2\frac{y}{z}\,dy,$$

and arranging this in the form

$$\frac{zdx - a^2y\,dy}{\sqrt{x^2 - a^2y^2}} = adx$$

we see that it is exact, and integrating we have

$$\sqrt{z^2 - a^2y^2} = ax + b.$$

The complete solution of the given equation is, therefore,

$$z^2 = a^2y^2 + (ax + b)^2.$$

EXERCISES LXIX

1. Find general solutions of the following linear differential equations of the first order.

(a) $p + q = z$; (d) $(y - z)p + (z - x)q = x - y$;
(b) $x^2p + y^2q = z^2$; (e) $yzp + xzq = y$;
(c) $yp - xq = x^2 - y^2$; (f) $y^2zp + x^2zq = xy^2$.

2. Find complete solutions of the following:

(a) $pq = 5$; (d) $p = 2xq^2$;
(b) $z = px + qy + p^2$; (e) $p = (z + yq)^2\,\phi$;
(c) $p(1 + q) = qz$; (f) $p(q^2 + 1) + (1 - z)q = 0$.

3. Find the general solution of the following:

(a) $xp^2 - 2zp + xy = 0$; (b) $(qy - z)^2 = x^2(q^2 + 1)$.

4. Solve the following:

(a) $y^2(p^2 - 1) = x^2p^2$; (c) $q^2 = z^2(p - q)$;
(b) $z = px + qy + (p + q)^2$; (d) $p + y(z - x) = 0$.

ART. 73. PARTIAL DIFFERENTIAL EQUATIONS OF HIGHER ORDER

Among partial differential equations of higher order than the first only a few types can be solved by elementary methods, and we must be content here to present those cases that permit of brief treatment. Fortunately many of the partial differential equations that arise in engineering and scientific problems yield to the methods of this brief discussion. While we shall remember the use of p, q, r, s and t to denote partial derivatives of first and second order, it will be profitable here to extend the notion of differential operator introduced in Chapter XV. We shall denote partial differentiation with respect to x by D_x and with respect to y by D_y; so that

$$D_xz = \frac{\partial z}{\partial x}, \quad D^2{}_yz = \frac{\partial^2 z}{\partial y^2}, \quad (D^3{}_x + 2D^2{}_y\,D_x)z = \frac{\partial^3 z}{\partial x^3} + 2\,\frac{\partial^3 z}{\partial y^2\,\partial x},$$

and so on, any polynomial in D_x and D_y indicating a definite operation which, when applied to z, is a linear expression in partial derivatives of z.

1. **Equations solved directly.** Frequently results may be obtained by treating the given equation as an ordinary differential equation, regarding one of the variables as constant. For example, take $r + p = xy$. This may be written

$$\frac{dp}{dx} + p = xy$$

if y is regarded as constant. The equation is linear of the first order, and we have

$$pe^x = y \int xe^x dx = ye^x(x - 1) + \phi_1(y),$$

therefore

$$p = \frac{dz}{dx} = (x - 1)y + \phi_1(y)e^{-x}.$$

Integrating again

$$z = \left(\frac{1}{2}x^2 - x\right)y - \phi_1(y)\, e^{-x} + \phi_2(y).$$

This is the general solution, involving two arbitrary functions.

As another example take $xs + q = xy^3$. We may write, after dividing by x

$$\frac{dq}{dx} + \frac{q}{x} = y^3$$

and solving for q, taking y as constant,

$$qx = \frac{1}{2}x^2y^3 + \phi(y).$$

Now integrating with respect to y, thinking of x as constant

$$zx = \frac{1}{8}x^2y^4 + \phi_1(y) + \phi_2(x),$$

since the integral of ϕ is an arbitrary function ϕ_1.

As a third example, consider $2r + s = y$, integrating with respect to x we have

$$2p + q = xy + \phi(y).$$

This is a partial differential equation, linear of the first order, and may be solved by the methods of the previous article. We may take

$$\frac{dx}{2} = \frac{dy}{1} = \frac{dz}{xy + \phi(y)}.$$

The first two members give $x - 2y = a$, and using this to replace x in the third member we have $(a + 2y)ydy + \phi(y)dy = dz$, or

$$z = \frac{1}{2}ay^2 + \frac{2}{3}y^3 + \phi_1(y) + b.$$

Solving for a and b we have

$$x - 2y = a,\ z - \frac{1}{2}(x - 2y)y^2 - \frac{2}{3}y^3 - \phi_1(y) = b$$

and the general solution may be written

$$z = \frac{1}{2}xy^2 + \phi_1(y) + \phi_2(x - 2y),$$

the terms in y alone being absorbed in the arbitrary function $\phi_1(y)$.

2. **Linear equations with constant coefficients.** An example of a linear partial differential equation with constant coefficients is

$$\frac{\partial^3 z}{\partial x^2\,\partial y} + 4\frac{\partial^2 z}{\partial y^2} + 3\frac{\partial^2 z}{\partial x\,\partial y} - \frac{\partial z}{\partial x} + 5z = \sin(x + y).$$

Derivatives of any order may appear, but each in a term by itself, to the first power, with a constant coefficient. The dependent variable z also may appear in one term, and the function on the right may be any function of x and y. Using the notion of differential operator, this may be written

$$(D_x^2 D_y + 4D_y^2 + 3D_xD_y - D_x + 5)z = \sin(x + y).$$

Equations of this type have properties similar to the ordinary differential equations treated in Arts. 66 and 67.

If $z = \phi_1(x, y)$ and $z = \phi_2(x, y)$ are solution, then $z = \phi_1(x, y) + \phi_2(x, y)$ is a solution. Also, to solve the equation we first equate the

left member to zero and solve, calling the resulting expression for z the **complementary function,** then find a **particular integral** which yields the right member when put for z in the left member. The sum of the complementary function and the particular integral is the **general solution.**

(a) **Equations homogeneous in the derivatives.** Here the differential operator is homogeneous in the D's, that is, z is absent and all derivatives are of the same order. Let the equation be

$$F(D_x, D_y)z = 0. \tag{1}$$

If we put $z = \phi(y + mx)$ the left member reduces to the derivative of ϕ of the same order as the differential operator, multiplied by a polynomial in m of the form F(m, 1). The equation

$$F(m, 1) = 0 \tag{2}$$

is called the **auxiliary equation,** and if its roots are $m_1, m_2, \ldots . m_n$, then the solution of (1) is

$$z = \phi_1(y + m_1x) + \phi_2(y + m_2x) + \ldots . + \phi_n(y + m_nx). \tag{3}$$

If the values of m are all **distinct** then (3) is the general solution, since the number of arbitrary functions is equal to the order of the given equation. In case of **repeated roots,** the corresponding terms of (3) are multiplied by powers of x to render them independent. Thus, if m_1, m_2 and m_3 are all equal the corresponding three terms of (3) become ϕ_1, $x\phi_2$, $x^2\phi_3$. In case of a pair of **complex roots** the two corresponding terms of (3) may be combined to form two arbitrary functions that are real in form. Thus if $m_1 = a + bi$, $m_2 = a - bi$, we may put $\phi_1 = \theta_1 + i\theta_2$, $\phi_2 = \theta_1 - i\theta_2$, where θ_1 and θ_2 are arbitrary functions, and the corresponding terms of (3) become

$$\theta_1(y + ax + ibx) + i\theta_2(y + ax + ibx) = \phi_1(y + m_1x)$$

$$\theta_1(y + ax - ibx) - i\theta_2(y + ax - ibx) = \phi_2(y + m_2x).$$

Evidently the imaginary parts of the upper terms are of opposite signs to those in the lower terms, so that the sum is real.

For example take

$$(D_x^3 - 4D_x^2Dy + 5D_xD_y^2 - 2D_y^3)z = 0.$$

Equation (2) becomes

$$m^3 - 4m^2 + 5m - 2 = 0,$$

and the roots are 1, 1 and 2. The general solution is

$$z = \phi_1(y + x) + x\phi_2(y + x) + \phi_3(y + 2x).$$

It should be clear that x and y may be interchanged in this discussion. In case the differential operator contains a higher power of D_y than of D_x this interchange may be profitable. However this may be avoided by careful attention to **infinite roots** of the auxiliary equation. Consider the equation

$$(D_x^3D_y^2 - 4D_xD_y^4)z = 0.$$

The equation is of the fifth order, and the auxiliary equation must be regarded as of the fifth degree. We write

$$0m^5 + 0m^4 + m^3 + 0m^2 - 4m + 0 = 0.$$

The five roots are

$$\infty, \infty, 2, -2, 0,$$

and the solution is

$$z = \phi_1(x) + y\phi_2(x) + \phi_3(y + 2x) + \phi_4(y - 2x) + \phi_5(y).$$

If we had interchanged x and y in the discussion we should have had for the auxiliary equation

$$0m^5 - 4m^4 + 0m^3 + m^2 + 0m + 0 = 0$$

and the roots would have been

$$\infty, \frac{1}{2}, -\frac{1}{2}, 0, 0,$$

giving for the solution

$$z = \phi_1(y) + \phi_2(x + \frac{1}{2}y) + \phi_3(x - \frac{1}{2}y) + \phi_4(x) + y\phi_5(x).$$

Evidently these two solutions are of like significance.

We have now to consider the **particular integral** in case the right member is not zero. Very frequently the use of a trial integral is effective. For example take

$$(D_x^2 - 3D_xD_y + 2D_y^2)z = e^x(1 + e^y).$$

The auxiliary equation is $m^2 - 3m + 2 = 0$, and the roots are 1, 2, giving the complementary function

$$\phi_1(y + x) + \phi_2(y + 2x).$$

Consider the particular integral in two parts, corresponding to the terms e^x and e^xe^y on the right. For the first try $z = ae^x$, and we find $a = 1$. For the second we observe that $e^xe^y = e^{x+y}$ and this is not independent of $\phi_1(y + x)$, therefore we may try $z = bxe^xe^y$ or $z = Cye^xe^y$. Using the former we find $b = -1$, and the general solution is

$$z = \phi_1(y + x) + \phi_2(y + 2x) + e^x - xe^xe^y.$$

If we use $z = Cye^xe^y$ we find $C = 1$, and we might replace the last term just written by ye^xe^y. If we replace $\phi_1(y + x)$ by $\phi_1(y + x) + (y + x)\,e^xe^y$, which is still an arbitrary function of $y + x$, the solution as written reduces to the new form, showing that the two forms are of like significance.

Again take

$$(D_x^3 + D_x^2D_y)z = \sin x \cos y + x^2 + y^2.$$

The complementary function is

$$\phi_1(y - x) + \phi_2(y) + x\phi_3(y).$$

For that part of the particular integral which is to yield x^2 we have only to consider $D_x^3z = x^2$ and the necessary term is $\frac{1}{60}x^5$. For y^2 we observe that y^n and xy^n are not independent of the complementary function, hence try $z = ax^2y^n$ and find $n = 3$, $a = \frac{1}{6}$. Next we note that

$$\sin x \cos y = \frac{1}{2}\sin (x + y) + \frac{1}{2}\sin (x - y),$$

and treat the two terms separately. Observing that only derivatives of odd order occur, we try

$$z = b \cos (x + y),$$

and this yields $\frac{1}{2}\sin (x + y)$ if $b = \frac{1}{4}$. Since an arbitrary function of

$y - x$ is included in the complementary function, we must combine either x or y with $\sin(x - y)$ and $\cos(x - y)$. By trial of

$$z = C_1 y \sin(x - y) + C_2 y \cos(x - y)$$

we find $C_1 = -\frac{1}{2}$, $C_2 = 0$. The general solution is, therefore,

$$z = \phi_1(y - x) + \phi_2(y) + x\phi_3(y) + \frac{1}{60}x^5 + \frac{1}{6}x^2y^3 + \frac{1}{4}\cos(x + y) - \frac{1}{2}y \sin(x - y).$$

If, by inspection and trial, we are not able to find the particular integral we may proceed by **factoring the differential operator,** thus forming a chain of equations of the first order. To illustrate we solve again the example just considered. It may be seen that

$$(D_x^3 + D_x^2 D_y)z = (D_x + D_y)(D_x)(D_x)z,$$

where on the right the three differential operators of the first order are to be applied successively. We may write

$$D_x z = u$$

$$D_x u = v$$

$$(D_x + D_y)v = \sin x \cos y + x^2 + y^2.$$

Treating first the term x^2, we may consider

$$p + q = x^2,$$

and the associated system

$$\frac{dx}{1} = \frac{dy}{1} = \frac{dv}{x^2}.$$

We are seeking a particular integral, therefore we use the simplest relation found from this system which involves v; namely, $v = \frac{1}{3}x^3$. Now we find at once $u = \frac{1}{12}x^4$ and $z = \frac{1}{60}x^5$, which agrees with the result previously obtained. For y^2 we obtain the system

$$\frac{dx}{1} = \frac{dy}{1} = \frac{dv}{y^2},$$

therefore $v = \frac{1}{3}y^3$, $u = \frac{1}{3}xy^3$, $z = \frac{1}{6}x^2y^3$.

Changing the product sin x cos y into the sum

$$\frac{1}{2}\sin(x+y)+\frac{1}{2}\sin(x-y)$$

as before, we reach for the former of these terms the system

$$\frac{dx}{1}=\frac{dy}{1}=\frac{dv}{\frac{1}{2}\sin(x+y)}=\frac{dx+dy}{2}.$$

We find from this $v=-\frac{1}{4}\cos(x+y)$, therefore $u=-\frac{1}{4}\sin(x+y)$ and $z=\frac{1}{4}\cos(x+y)$. Finally, for the remaining term we reach the system

$$\frac{dx}{1}=\frac{dy}{1}=\frac{dv}{\frac{1}{2}\sin(x-y)}.$$

From the first two members we have $x-y=C$, and using this to simplify the last two we have

$$\frac{dy}{1}=\frac{dv}{\frac{1}{2}\sin C}$$

or $v=\frac{1}{2}y\sin C=\frac{1}{2}y\sin(x-y)$, therefore

$$u=-\frac{1}{2}y\cos(x-y) \text{ and } z=-\frac{1}{2}y\sin(x-y).$$

(b) **Equations not homogeneous in derivatives.** In this case the differential operator is not homogeneous, and, in general, cannot be resolved into linear factors. If, however, it can be separated into two or more factors, linear or not, the problem of solving is somewhat simplified. Suppose

$$F(D_x, D_y) = F_1(D_x, D_y)\,F_2(D_x, D_y),$$

where the functions are regarded as algebraic functions of D_x and D_y and the factors on the right are polynomials of the same type as that on the left. Now any solution of either of the equations

$$F_1(D_x, D_y)z = 0, \quad F_2(D_x, D_y)z = 0,$$

is a solution of

$$F(D_x, D_y)z = 0,$$

and the sum of the solutions of the first two is a solution of the last equation. The idea, of course, extends to any number of factors.

If one factor of the differential operator is linear, say $D_x - m D_y - k$, then we find one term in the solution by solving the equation

$$p - mq = kz. \tag{4}$$

The general solution of this may be written

$$z = e^{kx}\phi(mx + y). \tag{5}$$

In this one or both of the constants m and k may be zero. In case of a factor which does not contain D_x, say $D_y + k$, we solve the equation

$$q = -kz \tag{6}$$

and obtain for a term of the solution

$$z = e^{ky}\phi(x). \tag{7}$$

After finding terms of the solution that arise from linear factors of the differential operator, let the remaining factor be $F_1(D_x, D_y)$, then we must seek additional terms of the solution as solutions of the equation

$$F_1(D_x, D_y)z = 0. \tag{8}$$

If we try $z = Ce^{ax+by}$ the left member reduces to

$$Ce^{ax+by} F_1(a, b),$$

and this will be zero if a and b satisfy the equation $F_1(a, b) = 0$. This is a polynomial in a and b, and if we solve for a in terms of b we shall find several values, say $f_1(b)$, $f_2(b)$, etc. A term in the solution is

$$z = Ce^{f_1(b)x+by}$$

where b and C are arbitrary, and we may form any number of such terms by giving b and C different sets of values and adding results. Thus we include in the solution in the terms

$$z = \Sigma Ce^{f_1(b)x+by} + \Sigma Ce^{f_2(b)x+by} + \ldots . \tag{9}$$

The solution (9) is valid for (8) whether the differential operator in (8) contains linear factors or not, but if linear factors are present the terms in form (5) and (7) are preferable to the summations in (9). Linear factors present will be revealed in solving the auxiliary equation

$$F_1(a, b) = 0, \tag{10}$$

a term of form (5) coming from a solution $a = mb + k$ and a term of form (7) from a solution $b = k$.

In case a factor $a - mb - k$ is repeated the corresponding terms of the solution may be written

$$e^{kx}\,\phi_1(mx + y),\ xe^{kx}\,\phi_2(mx + y),\ x^2e^{kx}\,\phi_3(mx + y), \text{ etc.}$$

If a factor $b - k$ is repeated the corresponding terms are

$$e^{ky}\,\phi_1(x),\ ye^{ky}\,\phi_2(x),\ y^2e^{ky}\,\phi_3(x), \text{ etc.}$$

If a factor $a - f(b)$ is repeated the corresponding parts of (9) are

$$\Sigma Ce^{f(b)x+by},\ x\Sigma Ce^{f(b)x+by},\ x^2\Sigma Ce^{f(b)x+by}, \text{ etc.},$$

the values of C and b in the different summations being entirely arbitrary.

To find the particular integral when the right member is not zero, we may use the method of inspection and the trial integral, the method not being essentially different from that illustrated for the homogeneous case above.

If the differential operator may be separated into two or more polynomial factors we may substitute a chain of equations of lower order. Say the given equation may be written

$$F_1(D_x, D_y) \cdot F_2(D_x, D_y)z = f(x, y),$$

then we have only to treat the two equations

$$F_2(D_x, D_y)z = u,$$

$$F_1(D_x, D_y)u = f(x, y).$$

Consider the example

$$(D_x^4 - D_x^2D_y^2 - 4D_x^2 + 4D_y^2)z = \sin x.$$

The auxiliary equation is

$$a^4 - a^2b^2 - 4a^2 + 4b^2 = 0.$$

The solutions of this are

$$a = 2, \quad a = -2, \quad a = b, \quad a = -b,$$

all distinct and each leading to form (5). The complementary function is

$$e^{2x}\,\phi_1(y) + e^{-2x}\,\phi_2(y) + \phi_3(x + y) + \phi_4(y - x).$$

To find the particular integral we try $z = C \sin x$, since all derivatives are of even order, and we find $C = -\frac{1}{3}$.

As another example take

$$(D_x^2\,D_y^2 + D_y^3)z = e^{3x}.$$

Here the auxiliary equation has the solutions

$$a = i\sqrt{b}, \quad a = -i\sqrt{b}, \quad b = 0, \quad b = 0.$$

The first two give terms of form (9), the last two of form (7), repeated, and the complementary function is

$$\Sigma c e^{by+i\sqrt{b}\,x} + \Sigma c e^{by-i\sqrt{b}\,x} + \phi_1(x) + y\,\phi_1(x).$$

For the particular integral we may try $z = Cy^2e^{3x}$ and find $C = \frac{1}{18}$, or we may write

$$(D_x^2\,D_y^2 + D_y^3)z = (D_x^2 + D_y)\,(D_y)\,(D_y)z = e^{3x}$$

and solve the system

$$D_y z = u,$$
$$D_y u = v,$$
$$(D_x^2 + D_y)v = e^{3x}.$$

We easily see that $v = \frac{1}{9}e^{3x}$, therefore $u = \frac{1}{9}ye^{3x}$ and $y = \frac{1}{18}y^2e^{3x}$.

EXERCISES LXX

1. Solve the partial differential equations:

(a) $yt - q = xy^3$; (c) $r + s + p = 0$;

(b) $s = xy$; (d) $xr = 4p$.

2. Solve the homogeneous equations:

(a) $(D_x^3 - 2D_x^2D_y - D_x D_y^2 + 2D_y^3)z = 0$;

(b) $(D_x^3 + 3D_x^2D_y)z = 0$;

(c) $(D_x^2D_y + 2D_xD_y^2 + D_y^3) = 0$.

3. Find the general solutions of the following:

(a) $(D_x^2 - D_xD_y)z = y + \sin x$;

(b) $(D_x^3 - 3D_xD_y^2 + 2D_y^3)z = e^{2x} + \cos(x - 2y)$.

4. Find the complementary functions of the following equations:

(a) $(D_xD_y^2 + D_y^4)z = \cos(x + y)$;

(b) $(D_x^2 + D_xD_y + 2D_x + D_y + 1)z = x^2 + y^2$;

(c) $(D_x^2 - 4D_x + 4)(D_y + 1)z = xy$.

5. Find the particular integrals of the equations in 4.

6. Solve $r - s + q - z = \cos(x + 2y) + e^y$.

7. The differential equation for torsional vibration of a shaft is

$$G\frac{\partial^2\theta}{\partial x^2} = m\frac{\partial^2\theta}{\partial t^2},$$

where θ is angular displacement of the shaft at any distance x from one end at the time t, G is the modulus of elasticity of the material of the shaft in shear, and m is the mass density of the material. Show that the general solution may be written in the form

$$\theta = \phi_1(\sqrt{m}\,x + \sqrt{G}\,t) + \phi_2(\sqrt{m}\,x - \sqrt{G}\,t).$$

8. Assuming that ϕ_1 and ϕ_2 above are harmonic functions of the same frequency, that is,

$$\phi_1(u) = a_1 \sin ku + b_1 \cos ku,$$

$$\phi_2(v) = a_2 \sin kv + b_2 \cos kv,$$

and that θ is zero for all values of t when $x = 0$ (i.e., the shaft is fixed at the origin of x), show that we have

$$\theta = (2a_1 \cos k\sqrt{G}\,t - 2b_1 \sin k\sqrt{G}\,t)\sin k\sqrt{m}\,x.$$

Show further that $b_1 = 0$ if the angular velocity of the shaft is zero when $t = 0$, at any distance x.

9. The differential equation for the transverse vibration of a bar or beam is

$$\frac{\partial^4 y}{\partial x^4} = -4h^4 \frac{\partial^2 y}{\partial t^2}$$

where h is given by the equation

$$4h^4 = \frac{mA}{EI},$$

m being the density of the material of the bar, A the cross-section area, E the modulus of elasticity in tension, and I the moment of inertia of the section area about the neutral axis. The dependent variable, y, is the deflection of the neutral axis at time t at distance x from some chosen origin. Find the solution in the form

$$y = \Sigma C e^{h\sqrt{b}(1+i)x+bt} + \Sigma C e^{h\sqrt{b}(1-i)x+bt}$$
$$+ \Sigma C e^{-h\sqrt{b}(1+i)x+bt} + \Sigma C e^{-h\sqrt{b}(1-i)x+bt}.$$

10. In the differential equation of 9 put $y = z \cos \omega t$, where z is a function of x only, and solve the resulting ordinary differential equation for z, obtaining for y the value

$$y = \cos \omega t(C_1 \sin kx + C_2 \cos kx + C_3 \sinh kx + C_4 \cosh kx),$$

where $k^4 = 4h^4\omega^2$. Putting these trigonometric and hyperbolic functions in exponential form show that each term of this product is a given by a term in the solution in 9 by assigning suitable values to b and C, real or complex.

11. Using the solution in 10 show that the maximum deflection for all points of the bar occurs when $t = 0$. If $C_2 = C_4 = 0$, show that a node (point of no deflection for all values of t) occurs at $x = 0$. If a second node occurs at $x = a$, show that $C_1 \sin ka = -C_3 \sinh ka$, or else that $C_3 = 0$ and $ka = 2n\pi$ where n is any integer. Discuss the form of the elastic curve of the bar at any time t under these assumptions.

BOOKS FOR SUPPLEMENTARY READING

Advanced Calculus, W. F. Osgood (Macmillan, 1925), Chapter XV.
Advanced Calculus, E. B. Wilson (Ginn, 1912), Chapter X.
Differential Equations, A. Cohen (Heath, 1906), Chapters XII–XIV.
Differential Equations, D. A. Murray (Longmans, 1919), Chapter XII.
Differential Equations, H. Bateman (Longmans, 1918), Chapters VI and VII.
Differential Equations, A. R. Forsyth (Macmillan, 1903), Chapters IX and X.

CHAPTER XVIII

DIFFERENTIAL EQUATIONS, ADDITIONAL DEVICES

Art. 74. Approximate Solution of Differential Equations

In the preceding chapters we have presented briefly the methods most commonly employed for solving differential equations. It is well understood, however, that many differential equations arise for which none of these methods are successful, and, in fact, for which no method is known by which the general solutions may be found. It should be as clearly understood that solutions always exist if the functions involved satisfy certain simple conditions as to continuity and as to possession of derivatives. It is the purpose of the present article to show how, in definite numerical cases with stated initial conditions, the solution may always be found as a tabulated function, the tabular values accurate to as many places as desired.

The numerical methods employed here bear a relation to the methods in previous chapters similar to the relation which the numerical substitutes for integration in Art. 45 bear to the calculus methods of integration. In fact, the trapezoidal rule and Simpson's one-third rule may be treated as special cases under the present article, just as ordinary integration is essentially the solution of a simple differential equation of the first order.

1. **Short arc integration, successive approximations.** Given a differential equation of the first order

$$p = \frac{dy}{dx} = f(x, y),$$

with the condition that $y = y_o$ when $x = x_o$, we may state that one and only one solution exists if the function on the right is a continuous function of x and has a partial derivative with respect to y for all values of x and y in the neighborhood of (x_o, y_o). Omitting critical discussion of this statement, we proceed to show how y may be tabulated as a function of x.

The initial value of p is $p_o = f(x_o, y_o)$, and an approximate value of y at $x_1 = x_o + h$ is $y_1^{(1)} = y_o + p_o h$. An approximate value of p at x_1 is $p_1^{(1)} = f(x_1, y_1^{(1)})$. A second approximation for y_1 is

$y_1^{(2)} = y_o + \frac{h}{2}(p_o + p_1^{(1)})$, where the rate of increase of y between x_o and x_1 is taken as the mean of the known p_o and the approximate $p_1^{(1)}$. Now a second approximation for p_1 is $p_1^{(2)} = f(x_1, y_1^{(2)})$. If h is not too large $p_1^{(2)}$ should be so near $p_1^{(1)}$ that the third approximation for y_1, $y_1^{(3)} = y_o + \frac{h}{2}(p_o + p_1^{(2)})$ will agree with $y_1^{(2)}$ to the number of places desired. In case there is still disagreement, however, the process may be carried on, using successively

$$p_1^{(k)} = f(x_1, y_1^{(k)}), \; y_1^{(k+1)} = y_o + \frac{h}{2}(p_o + p_1^{(k)})$$

until successive approximations are in essential agreement. If more than two or three successive approximations are required it will be found profitable to use a smaller value of h; the most advantageous value of h to use in a given instance can be known only by trial, although experience with examples of specific types may be an effective guide for further examples of similar types.

Supposing the successive values $y_1^{(1)}$, $y_1^{(2)}$, etc., have converged to a definite value y_1, we may now proceed to approximate y at $x_2 = x_1 + h$, by the same method, merely using x_1 and y_1 instead of x_o and y_o. The process may be carried on as long as the function f(x, y) possesses the properties essential for the existence of a solution.

To illustrate, let us take

$$p = \frac{dy}{dx} = -\frac{x}{y},$$

with the condition that $y = 5$ when $x = 0$. We have, then, $x_o = 0$, $y_o = 5$, $p_o = 0$. The following table shows the successive approximations for y and p at the chosen values of x. The value of h is taken as 0.5 at the start, but is reduced to 0.25 at $x = 1$ and to 0.1 at $x = 1.5$. All computations are carried to four decimal places.

$x_1 = 0.5$	$y_1^{(1)} = 5.0000$	$p_1^{(1)} = -0.1000$
	$y_1^{(2)} = 4.9750$	$p_1^{(2)} = -0.1005$
	$y_1^{(3)} = 4.9749$	$p_1^{(3)} = -0.1005$
$x_2 = 1.0$	$y_2^{(1)} = 4.9237$	$p_2^{(1)} = -0.2030$
	$y_2^{(2)} = 4.8990$	$p_2^{(2)} = -0.2041$
	$y_2^{(3)} = 4.8988$	$p_2^{(3)} = -0.2041$

$$
\begin{array}{lll}
x_3 = 1.25 & y_3^{(1)} = 4.8478 & p_3^{(1)} = -0.2578 \\
 & y_3^{(2)} = 4.8411 & p_3^{(2)} = -0.2582 \\
 & y_3^{(3)} = 4.8410 & p_3^{(3)} = -0.2582 \\
x_4 = 1.5 & y_4^{(1)} = 4.7764 & p_4^{(1)} = -0.3140 \\
 & y_4^{(2)} = 4.7695 & p_4^{(2)} = -0.3145 \\
 & y_4^{(3)} = 4.7694 & p_4^{(3)} = -0.3145 \\
x_5 = 1.6 & y_5^{(1)} = 4.7379 & p_5^{(1)} = -0.3377 \\
 & y_5^{(2)} = 4.7368 & p_5^{(2)} = -0.3377 \\
x_6 = 1.7 & y_6^{(1)} = 4.7030 & p_6^{(1)} = -0.3615 \\
 & y_6^{(2)} = 4.7018 & p_6^{(2)} = -0.3616 \\
x_7 = 1.8 & y_7^{(1)} = 4.6656 & p_7^{(1)} = -0.3858 \\
 & y_7^{(2)} = 4.6644 & p_7^{(2)} = -0.3859 \\
x_8 = 1.9 & y_8^{(1)} = 4.6258 & p_8^{(1)} = -0.4107 \\
 & y_8^{(2)} = 4.6246 & p_8^{(2)} = -0.4108 \\
x_9 = 2.0 & y_9^{(1)} = 4.5835 & p_9^{(1)} = -0.4363 \\
 & y_9^{(2)} = 4.5822 & p_9^{(2)} = -0.4364
\end{array}
$$

The example chosen here may be solved very easily, of course, by former methods, the solution corresponding to the given initial conditions being $x^2 + y^2 = 25$. When $x = 2$ we have $y = \sqrt{21} = 4.5826$, which shows that the method is very effective, the last digit retained being the only one in doubt.

Evidently the process may be carried on, but as x increases y will decrease more and more rapidly, and smaller and smaller values of h will be required to make the method effective. The method obviously fails at $x = 5$, for there p becomes infinite.

2. **Use of difference functions.** It is evident that we may make the approximations in the example above more rapidly by observing carefully the differences in values already found. For instance, in finding $y_4^{(1)}$ we took $y_3 + hp_3$, making no effort to predict the mean of p_3 and p_4. But observing that p_3 exceeds p_2 about -0.0541 we might estimate that p_4 will exceed p_3 by a similar amount and take the mean of p_3 and the estimated p_4. This would give the mean as about -0.2852 and we should have $y_4^{(1)} = 4.7697$, which is so near the cor-

rect value that the second approximation would be sufficient and the third approximation would be unnecessary.

In the present section we show how a systematic use of the differences of tabulated values already found may reduce the labor of computing further values.

Let t be the independent variable, and let u be tabulated as a function of t, the successive values of t having the common difference h. Suppose u and t entered in parallel columns up to t_n and u_n, with several preceding pairs of values as shown in the table. Let the a's represent the differences between successive u's, $a_k = u_k - u_{k-1}$, etc., and let the b's represent the differences between successive a's, called the second differences of the u's. Similarly let the c's represent the third differences, the d's the fourth differences, etc.

t	u	a	b	c	d
t_{n-4}	u_{n-4}				
t_{n-3}	u_{n-3}	a_{n-3}			
t_{n-2}	u_{n-2}	a_{n-2}	b_{n-2}		
t_{n-1}	u_{n-1}	a_{n-1}	b_{n-1}	c_{n-1}	
t_n	u_n	a_n	b_n	c_n	d_n

Now it is evident that the following equations result from the definitions of these differences:

$$u_{n-1} = u_n - a_n, \tag{1}$$

$$u_{n-2} = u_{n-1} - a_{n-1}, \tag{2}$$

$$u_{n-3} = u_{n-2} - a_{n-2}, \tag{3}$$

$$u_{n-4} = u_{n-3} - a_{n-3}, \tag{4}$$

$$a_n = a_{n-1} + b_n, \tag{5}$$

$$a_{n-1} = a_{n-2} + b_{n-1}, \tag{6}$$

$$a_{n-2} = a_{n-3} + b_{n-2}, \tag{7}$$

$$b_{n-1} = b_n - c_n, \tag{8}$$

$$b_{n-2} = b_{n-1} - c_{n-1}, \tag{9}$$

$$c_n = c_{n-1} + d_n. \tag{10}$$

If we add together equations (1), (2) and (5) we obtain

$$u_{n-2} = u_n - 2a_n + b_n, \tag{11}$$

and if we add to this result (3), (5), (6) and (8) we obtain

$$u_{n-3} = u_n - 3a_n + 3b_n - c_n. \tag{12}$$

If we add to (12) equations (4)–(10), using (8) doubled, we find

$$u_{n-4} = u_n - 4a_n + 6b_n - 4c_n + d_n. \tag{13}$$

This process may be used to find u_{n-r}, where r is any positive integer, in terms of u_n and the successive differences having the same subscript. Noting that the coefficients on the right of (11), (12) and (13) are the binomial coefficients, we may write

$$u_{n-r} = u_n - ra_n + \frac{r(r-1)}{1\cdot 2} b_n - \frac{r(r-1)(r-2)}{1\cdot 2\cdot 3} c_n + \ldots, \tag{14}$$

and this is true if r is a positive integer.

Now if we regard r in (14) as a continuous variable, and suppose the equation to terminate after $k + 1$ terms, then we have u given as a polynomial in r which gives the correct values of u for the values $r = 0, 1, 2, \ldots . k$. If we write $t = t_n - hr$ then (14) determines u as a polynomial function of t which is satisfied by the pairs of values tabulated, the number of pairs of values covered being equal to the number of terms retained in (14).

We may regard (14) as an **interpolation formula.** To approximate the value of u when t is midway between t_{n-1} and t_n, that is $t_n - \frac{1}{2}h$, we put $r = \frac{1}{2}$ in (14) and retain as many terms as necessary for the desired accuracy. In general, if (14) is carried to $k + 1$ terms and we evaluate for any fractional r between 0 and k we have the value of u at $t = t_n - hr$. If we evaluate (14) for r greater than k or negative, the process is called **extrapolation,** the value of u being estimated at a point beyond the range of known values. In particular, if we put $r = -1$ we obtain an estimate of u at $t = t_n + h$, namely

$$u_{n+1} = u_n + a_n + b_n + c_n + \ldots . \tag{15}$$

If we wish to integrate the function u with respect to t we observe that $dt = -h\,dr$, and therefore

$$\int_{t_n - hr_1}^{t_n - hr_2} u\,dt = -h \int_{r_1}^{r_2} u\,dr,$$

and using (14), multiplying out the product terms, we may easily evaluate the integral on the right. The result is only approximate, of course, since the polynomial (14) is only an approximation to the actual function u. The following particular case is important:

$$\int_{t_{n-1}}^{t_n} u\,dt = \int_{t_n-h}^{t_n} u\,dt = -h\int_1^0 u\,dr = h\int_0^1 u\,dr$$

$$= h(u_n - \frac{1}{2}a_n - \frac{1}{12}b_n - \frac{1}{24}c_n - \frac{19}{720}d_n \ldots\ldots). \quad (16)$$

Similarly we have

$$\int_{t_n}^{t_{n+1}} u\,dt = h(u_n + \frac{1}{2}a_n + \frac{5}{12}b_n + \frac{3}{8}c_n + \ldots\ldots). \quad (17)$$

By adding together (16) and (17) we obtain

$$\int_{t_{n-1}}^{t_{n+1}} u\,dt = h\left[2u_n + \frac{1}{3}(b_n + c_n + \ldots\ldots)\right]. \quad (18)$$

Let us apply these equations first to a problem in simple integration. In the theory of probability we find use for the expression

$$P = \frac{2}{\sqrt{\pi}}\int_0^t e^{-t^2}\,dt.$$

Let $u = e^{-t^2}$, and let us tabulate u and its first, second and third differences at intervals of one-tenth from $t = 0$ to $t = 1$. The values of u may be computed by series, or they may be taken from printed tables. In the column at the left, we may record values of

$$\frac{1}{2}\sqrt{\pi}\,P = \int_0^t u\,dt,$$

which may be computed by (16). First we have

$$\int_0^{0.1} u\,dt = 0.1\left[0.99005 - \frac{1}{2}(-0.00995)\right] = 0.09950.$$

Here $t_n = 0.1$, $t_{n-1} = 0$, $h = 0.1$, $u_n = 0.99005$, $a_n = -0.00995$.

Only the first difference is available at this first interval, but for the next we may use the second difference and in all other cases the approximation may be carried to third differences. Fourth differences

would contribute nothing to the results in this case. The increments in the left column are computed successively by (16). Thus the

$\frac{1}{2}\sqrt{\pi}\,P$	t	u	a	b	c
0.00000	0.0	1.00000			
0.09950	0.1	0.99005	−0.00995		
0.19720	0.2	0.96079	−0.02926	−0.01931	
0.29108	0.3	0.91393	−0.04686	−0.01760	0.00171
0.37950	0.4	0.85214	−0.06179	−0.01493	0.00267
0.46113	0.5	0.77880	−0.07334	−0.01155	0.00338
0.53500	0.6	0.69768	−0.08112	−0.00778	0.00377
0.60053	0.7	0.61263	−0.08505	−0.00393	0.00385
0.65751	0.8	0.52729	−0.08534	−0.00029	0.00364
0.70608	0.9	0.44486	−0.08243	0.00291	0.00320
0.74667	1.0	0.36788	−0.07698	0.00545	0.00254
		7.14607	−0.63212	−0.06703	0.02476

value at t = 0.5 is

$$0.37950 + \int_{0.4}^{0.5} u\,dt,$$

and by (16) the computation takes the form

value at t = 0.4	0.37950
1/10 of u at t = 0.5	77880
−1/20 of a at t = 0.5	3667
−1/120 of b at t = 0.5	96
−1/240 of c at t = 0.5	−14
value at t = 0.5	0.461129

Equation (17) is an alternative method of computing the same increment. Taking $t_n = 0.4$ the computation takes the form

value at t = 0.4	0.37950
1/10 of u at t = 0.4	85214
1/20 of a at t = 0.4	−3089
5/120 of b at t = 0.4	−620
9/240 of c at t = 0.4	100
value at t = 0.5	0.461105

This is not quite as accurate as the previous computation, since it involves "integration ahead," beyond the intervals over which differences are taken, but it supplies a check on the previous computation.

Again, we may compute two increments at once by (18). Taking $t_{n-1} = 0.3$ we have

value at t = 0.3	0.29108
2/10 of u at t = 0.4	170428
1/30 of b at t = 0.4	−477
1/30 of c at t = 0.4	89
value at t = 0.5	0.461120

Final checks against mistakes in computations are obtained by adding the columns for u, a, b, and c. The total in each column is the difference between the first and last entries in the preceding column. Also, if the first value of u is omitted from the total in the u column, then the total increment on the left is given by

$$\int_0^1 u\,dt = h\left[\Sigma u - \frac{1}{2}\Sigma a - \frac{1}{12}\Sigma b - \frac{1}{24}\Sigma c\right],$$

for this is the result of adding the ten equations by which the ten successive increments were computed when using (16). In the present case we have

1/10 of Σu	0.714605
−1/20 of Σa	31606
−1/120 of Σb	559
−1/240 of Σc	−103
Total	0.746667

The true value of P, as taken from a table of probability at t = 1, is 0.84270, and when this is multiplied by $\frac{1}{2}\sqrt{\pi}$, we find 0.74682, showing that our accumulated error is only 0.00015. Investigation shows that this is almost entirely due to omitting second and third differences at the start. If u had been tabulated at t = −0.1 and t = −0.2 so that these differences could have been used, then the results would have been in virtual agreement with five place tables for P.

Let us now see how the difference functions may be applied to the differential equation treated in the previous section. Suppose y_1 and

p_1 to have been computed by successive approximations as before, and construct the short table

x	y	p	a	b
0.0	5.0000	0.0000		
0.5	4.9749	−0.1005	−0.1005	

where a is the first difference for p. We may now apply (15) to estimate the value of p at x = 1.0, then form differences and compute y at x = 1.0 by (16). The new line in the table thus becomes

1.0	4.8996	−0.2010	−0.1005	0.0000

But when we put x = 1.0, y = 4.8996 in the differential equation we find p = −0.2041, and correcting the values for p, a and b, then computing again the value of y by (16) we find for the corrected third line of the table

1.0	4.8989	−0.2041	−0.1036	−0.0031.

If we compute p by the differential equation using this corrected value of y we shall find no change in p. Again extrapolating by (15) to find the next value of p, then forming differences and computing y by (16) we have for a fourth line of the table

1.5	4.7703	−0.3108	−0.1067	−0.0031

and computing p from this x and y we find p = −0.3144, leading to the corrected fourth line

1.5	4.7696	−0.3144	−0.1103	−0.0067.

We may proceed in this way to the values at x = 2.0, but to illustrate the method further we show a convenient adjustment to use shorter intervals. Accordingly we interpolate for p at x = 1.25, using (14) with u_n = p at x = 1.5 and taking $r = \frac{1}{2}$. The interpolated value of p is −0.2584, and now, for a new start, we have the short table written below. The value of y at x = 1.25 is computed by (16), taking x_n = 1.50, and subtracting the result from the value of y at x = 1.50.

x	y	p	a	b
1.00	4.8989	−0.2041		
1.25	4.8412	−0.2584	−0.0543	
1.50	4.7696	−0.3144	−0.0560	−0.0017.

Now applying (15) to estimate a new value of p and computing the new a and b, then applying (16) to find the new y we have for the approximate fourth line of the new table

1.75	4.6838	−0.3721	−0.0577	−0.0017,

but when we put these values of x and y in the differential equation we find $p = -0.3736$, and this gives for the corrected fourth line

1.75	4.6837	−0.3736	−0.0592	−0.0032.

We may now estimate the next value of b as about 45, and we readily construct the next line of the table

2.00	4.5824	−0.4373	−0.0637	−0.0045

and computing a new p with these values of x and y we find for the corrected line

2.00	4.5825	−0.4365	−0.0629	−0.0037

This final result is more nearly correct than that obtained in the previous section.

3. Numerical integration of systems of equations. In Art. 69 it was shown that a system of ordinary differential equations of any order may be reduced to a system of equations of the first order, and, in particular, a single equation of the nth order is reduced to a system of n equations of the first order. In solving differential equations numerically, therefore, it is sufficient to be able to treat systems of the first order.

We may present the essential features of the method of numerical integration by discussing a system of three equations of the first order, say

$$u = \frac{dx}{dt} = u(x, y, z, t),$$

$$v = \frac{dy}{dt} = v(x, y, z, t),$$

$$w = \frac{dz}{dt} = w(x, y, z, t).$$

Here u, v and w are used as function symbols on the right and as convenient symbols for the first derivatives on the left. Let the initial conditions be $x = x_0$, $y = y_0$ and $z = z_0$ when $t = 0$. From the

given equations, then, we may compute u_o, v_o and w_o, and taking a suitable value of h we have as first approximations

$$x_1^{(1)} = x_o + u_o h, \quad y_1^{(1)} = y_o + v_o h, \quad z_1^{(1)} = z_o + w_o h$$

and using these values of x, y and z and putting $t = h$ approximate values of u_1, v_1 and w_1 are found. Now first differences, $u_1 - u_o$, etc., are available and second approximations $x_1^{(2)}$, $y_1^{(2)}$ and $z_1^{(2)}$ are made, and putting these in the differential equation second approximations of u_1, v_1 and w_1 are computed. The process is carried on until no further corrections are found necessary, just as in the case of one dependent variable in the previous section. Next, we estimate u_2, v_2 and w_2 by extrapolation, using (15) as far as first differences. Computing first and second differences in lines of u_2, v_2 and w_2 we proceed to approximate x_2, y_2 and z_2 by use of (16). With these values of x_2, y_2, z_2, and $t = 2h$ we compute from the given equations our second approximations of u_2, v_2, w_2, correct our differences and compute corrected values of x_2, y_2, z_2. We continue until no further corrections are necessary. Evidently the process for three equations differs from that in one equation treated in the previous section only in the fact that three sets of computations must be carried on simultaneously, maintaining separate columns for t, x, u and successive differences of u, for y, v and successive differences of v and for z, w and successive differences of w.

Let us illustrate the method indicated by consideration of a compound pendulum. The equation of motion is

$$I \frac{d^2\theta}{dt^2} = - Wh \sin \theta.$$

Where I is the moment of inertia of mass about the axis of suspension, W is the weight and h is the distance from the axis of suspension to the center of gravity, while θ is the angular displacement from the vertical at time t. Suppose the equation to reduce to

$$\frac{d^2\theta}{dt^2} = - 10 \sin \theta,$$

and let $\theta = 0.5$ radians and $\frac{d\theta}{dt} = 0$ when $t = 0$. To reduce this to a system of the first order we write

$$\frac{d\theta}{dt} = u \tag{19}$$

$$w = \frac{du}{dt} = - 10 \sin \theta. \tag{20}$$

We shall have

$$u_n = u_{n-1} + \int_{t_{n-1}}^{t_n} w\,dt \tag{21}$$

$$\theta_n = \theta_{n-1} + \int_{t_{n-1}}^{t_n} u\,dt \tag{22}$$

and to complete the cycle we have $w_n = -10 \sin \theta_n$. We arrange columns for t, for w and its successive differences, for u and its successive differences and for θ. The successive differences of θ are not necessary, but they are useful in detecting errors, since errors are likely to cause irregularities in the differences of θ. The value of h is taken conveniently as 0.1. At t = 0.0 we have $\theta = 0.5000$, therefore w = −4.794, and u = 0.0. To start the line at t = 0.1 we take as an approximate value of w the same value that w has at t = 0.0, and the approximate u is this multiplied by h, namely −0.4794. This is also the approximate value of the first difference of u. Now applying (16) and (22) we find the increment in θ is −0.0240, and entering this in the column of first differences for θ and adding it to the initial θ we find $\theta = 0.4760$. Putting this value of θ in (20) we have the second approximation for w at t = 0.1, namely w = − 4.582, giving

t	w	a	b	c
0.0	−4.794			
	−4.794			
	−4.582	0.212		
0.1	−4.587	0.207		
	−4.380	0.207		
	−3.960	0.627	0.420	
0.2	−3.966	0.621	0.414	
	−2.931	1.035	0.414	
	−3.046	0.920	0.299	−0.115
0.3	−3.044	0.922	0.301	−0.113
	−1.934	1.110	0.188	−0.113
	−1.598	1.446	0.524	0.223
0.4	−1.603	1.441	0.519	0.218
	0.857	2.460	1.019	0.500
	−0.077	1.526	0.085	−0.434
0.5	−0.064	1.539	0.098	−0.421

t	u	a	b	c
0.0	0.0000			
	−0.4794	−0.4794		
	−0.4688	−0.4688		
0.1	−0.4690	−0.4690		
	−0.9173	−0.4483	0.0207	
	−0.8999	−0.4309	0.0381	
0.2	−0.9001	−0.4311	0.0379	
	−1.2485	−0.3484	0.0827	0.0448
	−1.2527	−0.3526	0.0785	0.0406
0.3	−1.2526	−0.3525	0.0786	0.0407
	−1.5026	−0.2500	0.1025	0.0239
	−1.4918	−0.2392	0.1133	0.0347
0.4	−1.4901	−0.2375	0.1150	0.0364
	−1.5380	−0.0479	0.1896	0.0746
	−1.5730	−0.0829	0.1546	0.0396
0.5	−1.5724	−0.0823	0.1552	0.0402

t	θ	a	b	c
0.0	0.5000			
	0.4760	−0.0240		
	0.4766	−0.0234		
0.1	0.4766	−0.0234		
	0.4071	−0.0695		
	0.4078	−0.0688		
0.2	0.4078	−0.0688	−0.0454	
	0.2995	−0.1083		
	0.2993	−0.1085		
0.3	0.2993	−0.1085	−0.0397	0.0057
	0.1605	−0.1388		
	0.1610	−0.1383		
0.4	0.1610	−0.1383	−0.0298	0.0099
	0.0077	−0.1533		
	0.0064	−0.1546		
0.5	0.0064	−0.1546	−0.0163	0.0135

0.212 as first difference. Now applying (16) with (21) and (22) we complete the second approximation for the line at t = 0.1, obtaining θ = 0.4766. Using this θ in (20) we find w = −4.587, and thus start a third approximation, leading again to the same value of θ.

To start the line at t = 0.2 we assume the same increment in w as occurred from t = 0.0 to t = 0.1, namely 0.207. Entering this in the a column and adding to the previous w we have w = −4.380, and (16), (21) give −0.4483 as the increment in u. We enter this as first difference, compute u and b, then from (16) and (22) find −0.0695 as increment in θ, whence θ = 0.4071. This θ yields w = −3.960, which leads ultimately to θ = 0.4078, and the next round leads back to this same θ.

As we proceed higher differences are available, and with h = 0.1 third differences are quite important. With smaller values of h third differences might be negligible. Also, with smaller values of h the number of successive approximations for each line of the table might be reduced from three to two, or even to one, after the work has been carried far enough to obtain higher differences.

The method employed in the computation above for starting a new line is essentially extrapolation by (15) to find the next w, but in practice it consists in assuming that the highest difference available is the same in the new line then working backward to w. For example we have c = 0.113 at t = 0.3 and c =0.218 at t = 0.4, and this leads to the rough estimate that c = 0.500 at t = 0.5, and from this the rest of the line is readily computed, u and θ being obtained by "integrating across" by (16). It is ultimately found that c = −0.421 instead of 0.500, the significant fact being that an error of 0.900 in the first estimate of w is entirely corrected in the third approximation.

Another method for starting a new line would be to "integrate ahead" by (17). If we integrate u to find θ, taking t_n = 0.4 we find the increment in θ to be −1.547, giving θ = 0.0063. Computing w from this value of θ and integrating across by (16) we find θ = 0.0064, as previously.

Also, we may integrate ahead by (18). Again taking t_n = 0.4 and integrating u we find the increment in θ from t = 0.3 to t = 0.5 to be −0.2930, giving at t = 0.5 the value θ = 0.0063 as before.

It is evident that in the present instance, after higher differences are available, the method of integration ahead is superior to the method of extrapolation. This is due to the irregular character of third differences of w with the value of h used. The choice of method for starting the new line in any particular instance should be governed by the characteristics revealed by the computations as they progress.

The example treated here is typical of the systems of differential equations which arise from single equations of higher order, the computations coming in cycles, w, u, θ, w u, etc., the cycle being repeated until no corrections are found. In such case, an estimated value of any one of the quantities θ, u or w may be sufficient to start a new line of the table. In general, however, the functions corresponding to the right members of (19) and (20) may contain more than one of the quantities, and it may be necessary to estimate new values for several of the quantities to start a new line. These estimates may be made either by integration ahead or by extrapolation.

EXERCISES LXXI

1. Given the differential equation $\frac{dy}{dx} = \log_{10} \frac{x}{y}$, what values of x and y must be avoided in stating initial conditions? Assuming $y = 10$ when $x = 5$, tabulate y at intervals of 0.1 from $x = 5$ to $x = 6$.

2. Solve problem 10 of Exercises XLII by use of difference functions.

3. Tabulate $\int_{\frac{\pi}{6}}^{x} \log_{10} \sin x \, dx$ at intervals of one degree from $x = 30°$ to $x = 40°$.

4. Given $\frac{d^2s}{dt^2} = \log_{10} \cos s$, and $s = 0$, $v = 1$, at $t = 0$, construct the solution, choosing a suitable value of h. At what values of s will the method fail?

5. Construct the solution of

$$\frac{d^2\theta}{dt^2} + \frac{d\theta}{dt} = -\sin \theta,$$

given $\theta = 1$ and $\frac{d\theta}{dt} = 0$ when $t = 0$.

6. Tabulate x and y as functions of t, given

$$\frac{dx}{dt} = y + t, \quad \frac{dy}{dt} = e^x$$

subject to the condition $x = 0$, $y = 1$ when $t = 0$.

7. Tabulate x, y and z as functions of t, given

$$\frac{dx}{dt} = y - z, \quad \frac{dy}{dt} = z - x, \quad \frac{dz}{dt} = \sin (x - y),$$

with $x = 0$, $y = 0$, $z = 1$, when $t = 0$.

Art. 75. Solution by Series

For the numerical methods of the previous article differential equations were reduced to systems of the first order. For the analytical methods of the present article it is desirable to consider all dependent variables but one eliminated, so that we deal with a single differential equation of any order. Such an equation may be written

$$\frac{d^n y}{dx^n} = F\left(x, y, \frac{dy}{dx}, \ldots, \frac{d^{n-1}y}{dx^{n-1}}\right), \tag{1}$$

if we may suppose the equation solved for the highest derivative. We shall give two methods for expressing y as a series in powers of $x - x_0$, the coefficients depending on the values assumed for y and its successive derivatives up to the $(n - 1)$ st at $x = x_0$.

1. **Use of Taylor's expansion.** Assume that y may be expressed as a power series

$$y = c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + c_3(x - x_0)^3 + \ldots, \tag{2}$$

then by Taylor's theorem we must have

$$c_0 = y_0, c_1 = \frac{dy_0}{dx_0}, \ldots c_k = \frac{1}{k!}\frac{d^k y_0}{dx_0^k}, \ldots,$$

where the zero subscripts indicate that the functions are evaluated at $x = x_0$. The values of these up to $k = n - 1$ may be assumed arbitrarily, being the n arbitrary constants of the general solution, and substituting these values in (1) we obtain c_n. To obtain c_{n+1} we differentiate (1) with respect to x, finding an expression for the $(n + 1)$ st derivative in terms of the nth and lower derivatives, which are known. The process may be carried on indefinitely, provided that all the successive derivatives of F may be calculated at the assumed initial values.

As a simple example let us treat the equation

$$\frac{dy}{dx} = -\frac{x}{y}$$

which was solved numerically in the previous article. Assume $y = y_0$ when $x = x_0$, then

$$\frac{dy_0}{dx_0} = -\frac{x_0}{y_0}, \quad \frac{d^2y_0}{dx_0^2} = -\frac{x_0^2 + y_0^2}{y_0^3},$$

$$\frac{d^3y_0}{dx_0^3} = -\frac{3(x_0^3 + x_0y_0^2)}{y_0^5}, \quad \frac{d^4y_0}{dx_0^4} = -\frac{3(5x_0^4 + 6x_0^2y_0^2 + y_0^4)}{y_0^7}.$$

We write, then

$$y = y_o - \frac{x_o}{y_o}(x - x_o) - \frac{1}{2}\frac{x_o^2 + y_o^2}{y_o^3}(x - x_o)^2$$

$$- \frac{1}{2}\frac{x_o^3 + x_o y_o^2}{y_o^3}(x - x_o)^3 - \frac{1}{8}\frac{5x_o^4 + 6x_o^2 y_o^2 + y_o^4}{y_o^7}(x - x_o)^4 + \ldots .$$

Here y_o is the arbitrary constant. If we put $x_o = 0$, $y_o = 5$, as in the previous article, we have

$$y = 5 - \frac{1}{10}x^2 - \frac{1}{1000}x^4 + \ldots .$$

As another example let us consider the equation

$$\frac{d^2\theta}{dt^2} = -10 \sin \theta$$

which was treated in the last section of the previous article. Let $\theta = \theta_o$ and $\frac{d\theta}{dt} = u_o$ when $t = 0$, then using the given equation and

$$\frac{d^3\theta}{dt^3} = -10 \cos \theta \frac{d\theta}{dt},$$

$$\frac{d^4\theta}{dt^4} = -10 \sin \theta \left(\frac{d\theta}{dt}\right)^2 - 10 \cos \theta \frac{d^2\theta}{dt^2},$$

$$\frac{d^5\theta}{dt^5} = -10 \cos \theta \left(\frac{d\theta}{dt}\right)^3 - 20 \sin \theta \frac{d\theta}{dt}\frac{d^2\theta}{dt^2}$$

$$+ 10 \sin \theta \frac{d\theta}{dt}\frac{d^2\theta}{dt^2} - 10 \cos \theta \frac{d^3\theta}{dt^3},$$

and others that may be derived by further differentiation, we may write for the coefficients

$$c_o = \theta_o, c_1, = u_o, c_2, = -5 \sin \theta_o,$$

$$c_3 = -\frac{5}{3} u_o \cos \theta_o, c_4 = -\frac{1}{12} \sin \theta_o (u_o^2 - 10 \cos \theta_o),$$

$$c_5 = \frac{1}{12} u_o (10 - u_o^2 \cos \theta_o), \text{ etc.}$$

These involve the two arbitrary constants θ_o and u_o, become numerical coefficients when these are assigned definite numerical values, and determine θ as a power series in t.

2. **Equating coefficients.** In some cases it may be convenient to take y in the form (2) and compute the necessary derivatives of y from (2) and substitute in (1) and then determine the c's by equating coefficients of like powers of x. For this purpose it is not necessary that (1) be solved for the highest derivative, and usually it will be desirable to put $x' = x - x_o$, so that the series is in powers of the new variable x'.

Considering again the simple equation

$$\frac{dy}{dx} = -\frac{x}{y},$$

we may write this

$$y\frac{dy}{dx} = -x,$$

and assuming (2) with $x_o = 0$, and taking the first derivative, we have

$$(c_o + c_1x + c_2x^2 + \ldots)(c_1 + 2c_2x + 3c_3x^2 + \ldots) = -x$$

Multiplying the two series together we have

$$-x = c_oc_1 + (c_1^2 + 2c_oc_2)x + (3c_1c_2 + 3c_oc_3)x^2$$

$$+ (4c_1c_3 + 2c_2^2 + 4c_oc_4)x^3 + (5c_1c_4 + 5c_2c_3 + 5c_oc_4)x^4 + \ldots$$

This yields the equations

$$c_oc_1 = 0,\ c_1^2 + 2c_oc_2 = -1,\ c_1c_2 + c_oc_3 = 0,$$

$$2c_1c_3 + c_2^2 + 2c_oc_4 = 0,\ c_1c_5 + c_2c_3 + c_oc_5 = 0, \text{ etc.}$$

From the first $c_o = 0$ or $c_1 = 0$. The second equation makes c_1 imaginary if $c_o = 0$, therefore we take $c_1 = 0$. The second equation now gives $c_2 = \frac{1}{2c_o}$, the third equation gives $c_3 = 0$ and from the fourth $c_4 = -\frac{1}{8c_o^3}$ and from the fifth $c_5 = 0$. We find then,

$$y = c_o - \frac{1}{2c_o}x^2 - \frac{1}{8c_o^3}x^4 + \ldots$$

If we put $c_o = 5$ this gives the same particular solution we found before

As a final example consider the equation

$$\frac{d^2y}{dx^2} + x^2\frac{dy}{dx} + (1 - x^2)y = \sin x.$$

Assume a form for y and compute the terms on the left:

$$y = c_o + c_1x + c_2x^2 + c_3x^3 + \ldots,$$

$$-x^2y = - c_ox^2 - c_1x^3 + \ldots,$$

$$x^2\frac{dy}{dx} = c_1x^2 + 2c_2x^3 + \ldots,$$

$$\frac{d^2y}{dx^2} = 2c_2 + 6c_3x + 12c_4x^2 + 30c_5x^3 + \ldots.$$

Also we have for the right member

$$\sin x = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 + \ldots.$$

Equating coefficients and solving we obtain

$$y = c_o + c_1x - \frac{1}{2}c_ox^2 - \frac{1}{6}(c_1 - 1)x^3 + \frac{1}{24}(3c_o - 2c_1)x^4 + \frac{1}{120}(7c_1 + 6c_o - 2)x^5 + \ldots.$$

This solution contains the two arbitrary constants c_o and c_1, and is the general solution of the given equation.

In more extensive works on differential equations various modifications of the methods presented here are shown, which help to avoid or surmount the difficulties that arise, and which extend the range of types of equations that may be successfully solved by series. Let it suffice here to conclude with a few brief remarks.

After a few terms the coefficients may all be zero, giving y as a polynomial in x. In such case the equation may usually be solved by other means.

It may be possible to discover a form by which the general term of the series can be written. In this case it is often possible to recognize the series as the Taylor expansion of a known function, so that the solution is given in finite form.

If the series does not terminate and if the law of the coefficients is not discovered, the first few terms of the series may be regarded as a practical means of obtaining approximate numerical values. The utility of the method will depend on the rapidity of convergence of the series, and this will depend on the choice of x_o, y_o, $\frac{dy_o}{dx_o}$, etc., and on

the magnitude of $x - x_0$. It is necessary to avoid initial values and values of x at which y or any of its successive derivatives become infinite, and these usually reveal themselves directly, or by inconsistent conditions for determination of coefficients, or by failure of the series to converge. While, without knowledge of the general term, it is not possible to test the series for convergence, obvious inability to secure a sufficiently close numerical approximation is an effective warning in practice.

In some cases where the method of Section 1 fails because of the indeterminate character of the successive derivatives at the initial values chosen, the method of section 2 may yield a solution involving a smaller number of constants than required for the general solution. A solution thus found may often be used as an aid in finding the general solution by a method somewhat analogous to "variation of parameters" in the solution of linear equations with constant coefficients. The method is suggested in exercise 5 below.

EXERCISES LXXII

1. Find the general solutions of the following as power series in x.

(a) $\dfrac{dy}{dx} = x + y^2$; (b) $\dfrac{dy}{dx} = \dfrac{2y}{1 - x^2}$.

2. (a) Given $x\dfrac{dy}{dx} = x^2 + y^2$ show that if y is expressed as a power series in x then no arbitrary constant can enter the coefficients.

(b) Given $(x + h)\dfrac{dy}{dx} = (x + h)^2 + y^2$, find y as a power series in x, assuming $y = y_0$ when $x = 0$. Show that if $h = 0$, then $y_0 = 0$, and compare with the result in (a).

3. Find the general solution of the following as a power series in x.

$$\frac{d^2y}{dx^2} - xy = \sin x$$

4. Solve the equation $\dfrac{d^2y}{dx^2} + x^2y = 0$, subject to the condition $y = 2$ and $\dfrac{dy}{dx} = 5$ when $x = 0$.

5. By the method of section 2 above show that the equation

$$F(D)y = x^2\frac{d^2y}{dx^2} - (x^2 + 4x)\frac{dy}{dx} + 4y = 0$$

has a solution in the form $y_1 = cx^4e^x$. To find the general solution try $y = vy_1 + w$, where v and w are functions of x to be determined by putting this value of y in the given equation. Show that the equation becomes

$$vF(D)y_1 + F(D)w + y_1\left[x^2\frac{d^2v}{dx^2} - (x^2 + 4x)\frac{dv}{dx}\right] + 2x^2\frac{dv}{dx}\frac{dy_1}{dx} = 0.$$

The first term is zero, since y_1 is a solution of the given equation. Choose v so that the expression in brackets is zero, then solve the resulting equation for w. Since v contains two arbitrary constants, any particular value of w will furnish the general solution of the given equation.

6. Solve, by the method suggested in 5, the equation

$$x\frac{d^2y}{dx^2} + \frac{dy}{dx} + y = 0.$$

BOOKS FOR SUPPLEMENTARY READING

Differential Equations, A. Cohen (Heath, 1906), Chapter XI.
Differential Equations, H. Bateman (Longmans, 1918), Chapters IX and XI.
Differential Equations, A. R. Forsyth (Macmillan, 1903), Chapters V and VI.
New Methods in Exterior Ballistics, F. R. Moulton (University of Chicago Press, 1926), especially Chapters III and V.
The Calculus of Observations, Whitaker and Robinson (Blackie and Son, 1924), Chapter XIV.
Interpolation, J. F. Steffensen (Williams & Wilkins, 1927).

INDEX

(Page numbers without parentheses locate discussions in the text, page numbers in parentheses locate references in the exercises)

Abscissa, 5
Absolute value, 22
—— ——, complex number, 291
Acceleration, (44), (52), (69), (192)
Addition and subtraction, series, 283
—— —— ——, complex numbers, 292
—— —— ——, simple harmonic functions, 305
Adiabatic expansion, 381, (487)
Algebraic equation, 88
—— functions, definition, 4
—— ——, graphs, 6
—— ——, derivatives, 49
—— —— of complex variable, 297
Amplitude, complex numbers, 291
——, simple harmonic functions, 303
Angle between lines in plane, 160
—— —— curves in plane, 195
—— —— lines in space, 330
—— —— planes, 333
Approximate integration, 249, 514
—— solution of equations, 115
—— —— —— differential equations, 514
Area, interpretation of integral, 73
—— of polygon, 150
—— by integration, plane, 220
—— —— ——, surface of revolution, 223
—— —— iterated integral, 244
——, numerical approximation, 249
—— of curved surfaces, 371
Argument, complex number, 291
Asymptotes, 10, 139, 153, 162
——, hyperbola, 176
——, general, 363
Auxiliary variable, 18
—— circles, ellipse, 174
—— equation, 449, 504
Average deviation, 411

Axis, of parabola, 172
——, major, minor, 173
——, transverse, conjugate, 176
——, neutral, 239
——, of reals, of imaginaries, 292

Base, natural or Naperian, 11
——, general, 59
——, common, 60
Beams, 6, (86), (209), (226), 237, 238, 243, (256), (513)
Belt friction, 11, (115), (491)
Bending moment, 6, 237, 238, 243
Bernoulli's equation, 439
Binomial equations, 93, 296
—— theorem, positive integers, 90
—— ——, general, (280)
Boundary values, 493
Bounds, roots of equations, 93
——, of a locus, 153

Cardioid, (220), (226)
Cartesian coordinates, 148
Catenary, (124), (209), (219)
Center of curvature, 214
—— —— gravity, 228
—— —— pressure, 238
Circle, as a conic section, 170
—— of curvature, 213
——, osculating, 210
Cissoid, (364)
Clairaut's equation, 444
Column, Gordon's formula, (407)
——, eccentrically loaded, (463)
Combinations, selections, 90
Complementary function, 455, 504
Complete primitive, 429
—— solution, 493
Complex quantities, 290
—— variable, 297

Computation rules, precision, 415
Conditional convergence, 277
Confocal conics, 185
Conic sections, 168
Conjugate axis, 176
—— complex numbers, 118
—— hyperbola, 177
—— point, 358
Conservative force field, 382
Consistency of equations, 102
Constant of integration, 71
Contact point, 358
Continuity, 34
Continuous variable, 3
Convergence of series, 274
—— —— ——, comparison test, 275
—— —— ——, test-ratio test, 275
—— —— ——, absolute, 276
—— —— ——, conditional, 277
—— —— ——, test by integration, 277
Curvature, 212
Curve tracing, 197
——, length of, 217
——, in space, 328
Cusp, 205, 358
Cycloid, (215), (219), (226), (233), (364)

Damped vibrations, 452
Definite integral, 79
Degree of polynomial, 88
—— of differential equation, 428
DeMoivre's theorem, 295
Density, 233
Dependent variable, 2
Depressed equation, 93
Derivatives, definition, 39
——, general rules, 45
——, algebraic functions, 49
——, trigonometric functions, 53
——, exponential and logarithmic functions, 58
——, implicit functions, 62
——, parametric equations, 63
——, successive, 66
——, uses of, 192
——, of series, 285
——, partial, 337
——, total, 337
Derived function, 42
Descartes' rule of signs, 118
Determinants, 96
Difference functions, 516
Differential, definition, 82
—— equations, first order, 427
—— ——, higher degree, 441
—— ——, higher order, 449
—— ——, several variables, 469
—— ——, total, 479
—— ——, partial, 490
—— ——, numerical integration, 514
—— ——, solution by series, 529
Differentiation (see Derivatives), definition, 42
Direction cosines, 330
Directrix, conic sections, 170
Discrete variable, 3
Discriminant, quadratic, 94
——, general, 201
Distance, between points, 149
——, point from line, 161
——, point from plane, 334
Divergence of series, 274
Division, synthetic, 92
——, complex numbers, 294
Dominating series, 281
Double roots, 201
—— law of the mean, 261
—— point, 358
—— integrals, 366

Eccentricity of conic section, 170
Electric circuit, (115), (123), (310), (326), (441), (454), (464), (478)
Eliminant of set of equations, 104
Elimination, 96
——, Sylvester's method, 105
Ellipse, 172
Ellipsoid, 335
Empirical functions, definition, 4
—— data, treatment of, 385
—— equations, 418
Energy, kinetic, 236
——, intrinsic, 380
——, potential, 382
Entropy, 380, (488)
Envelopes, 358
Equation of the second degree, 179
Errors of observation, 388

Errors, residual, 395
——, persistent, 409
——, accidental, 410
——, mean, 411
——, percentage, 415
Euler's expressions, sine and cosine, 298
Evolutes, 361
Evolution, complex numbers, 295
Exact differential, 376
—— —— equation, first order, 436
—— —— ——, higher order, 466
Expansion of determinants, 98
—— in power series, 272
—— of gas, isothermal, 380
—— —— ——, adiabatic, 381
Exponential equations, 114
—— functions, graphs, 11
—— ——, derivatives, 58
—— —— of complex variable, 297
—— ——, semi-logarithmic plotting, 421
Extrapolation, 518

Factor theorem, 88
Factorial notation, 89
Falling body, 7, (52)
Family of curves, 155, 184
Finite discontinuity, 35
Focus, conic section, 170
Force field, 381, (476), (488)
Fourier series, 311
Frequency, 303
Friction, belt, 11, (115)
——, thrust bearing, 85
——, sliding, (113)
——, pivot, 239
Functions, definition, 1
——, gamma, 143
——, complex variable, 297
——, several variables, 327
——, implicit, derivatives, 62
——, ——, partial and total derivatives, 342

Gamma function, 143
Gas, perfect, 379
General solution, differential equation, 429
General solution, partial differential equation, 493
Gordon's formula, columns, (407)
Graphical multiplications and division 17
—— elimination, 17
—— solution of equations, one unknown, 116
—— —— —— ——, two unknowns, 121
—— treatment of complex quantities, 291
—— representation, functions of complex variable, 300
Graphs of simple functions, 5
—— —— compound functions, 15
Green's theorem, 375
Gudermannian function, (20), (61)

Harmonic series, (279)
—— functions, simple, 303
—— analysis, 311
—— motion, 451
Helicoidal surface, 336
Helix, 336
Homogeneous differential equations, first order, 432
—— —— ——, linear, 449
—— total differential equations, 482
Hyperbola, 175
Hyperbolic functions, derivatives, 61
—— —— of complex variable, 299
Hyperboloid, 335

Implicit functions, derivatives, 62
—— —— partial and total derivatives, 342
Impulse, 235
Increment, 40
Independence of equations, 102
Independent variable, 2
Indeterminate forms, 258
Infinite discontinuity, 35
—— series, 273
—— limits of integration, 141
Infinity, ideal element, 29
Inflexion, point of, 198
Initial conditions, 71
—— ——, differential equation, 430
—— line, polar coordinates, 149

Integrability, total differential equations, 480
Integral, notation, 70
——, particular, 71, 455, 505
——, definite, 79
——, indefinite, 80
—— curve, 72, 434
—— ——, approximate construction, 252
——, probability, 143, 392
——, elliptic, 143
——, iterated, 244
——, multiple, 365
Integrand, 70
—— continuous, 137
—— discontinuous, 138
Integrating factor, 436
Integration, meaning, 70
——, standard forms, 74
——, by substitution, 125
——, by parts, 128
——, by trial integral, 130
——, rational fractions, 133
——, discontinuous integrand, 138
——, infinite limits, 141
——, use of tables, 145
——, applications, 217
—— of series, 284
Intercepts, 157
Intermediate value theorem, 258
Interpolation formula, 518
Intrinsic energy, 380
Inverse trigonometric equations, 112
Inversion of order, 97
Involution, complex numbers, 295
Isolated point, 380
Isothermal explansion, 380
Iterated integrals, 244

Kinetic energy, 236

Latus rectum, parabola, 172
—— ——, ellipse, 174
—— ——, hyperbola, 176
Law of the mean, 258
Least squares, 395
Legendre's linear equation, 461
Lemniscate, (196)
Length, tangent and normal lines, 193
Length, polar tangent and normal lines, 195
——, tangent to circle, 186
—— of arc of curve, 217
—— of diagonal of parallelopiped, 330
Limits of functions, 21
—— —— integration, 80
Line, standard equations, 157
——, tangent and normal, 187, 192
—— in space, 329
—— tangent to space curve, 348
—— normal to surface, 347
—— integrals, 371
Linear differential equations, first order, 438
—— —— ——, constant coefficients, 449
—— —— ——, systems of, 472
—— —— ——, partial, first order, 495
—— —— ——, ——, constant coefficients, 503
Locus of an equation, 151
Logarithms, graphs, 12
——, differentiation, 58
—— of complex quantities, 298
——, numerical precision, 415
Logarithmic plotting, 420
—— spiral, (215), (220)

Mass, 233
Maxima and minima, 203
—— —— ——, several variables, 351
Mean values, 227, 369
—— error, 411
—— value theorem, 259
Measure of precision, 392
Meridians, 335
Method of least squares, 393
Multiplication, graphical, 17
—— of roots of equation, 91
—— of series, 283
—— of complex numbers, 293
——, contracted method, 416
Mid-ordinate rule, 250
Mistakes in observations, 410
Modulus, logarithms, 60
——, complex numbers, 291
Moments, 227
——, statical, 228

Moments of inertia, 228
—— by multiple integration, 369
Momentum, 235
Moving bodies, (81), (82), 235, (440), (453), (463), (478)
Multiple roots, 201
—— integrals, 365

Neutral axis, 239
Normal equation of line, 158
—— —— of plane, 331
—— equations, observations, 401
—— lines, plane curves, 187, 192
—— ——, surfaces, 347
—— planes, space curves, 348
Numerical integration, 249, 514
—— ——, systems of equations, 523
—— calculations, precision, 411

Observation equations, linear, 400
—— ——, general, 404
Observations, direct, 393
——, ——, weighted, 398
——, indirect, 400
Octants, 328
Order, inversions of, 97
—— of determinant, 98
—— of differential equation, 427
—— of derivative, 67
Ordinary differential equations, 427
Ordinate, 5
Orifice, flow through, 86
Orthogonal systems of circles, 186
—— trajectories, 434
Osculating circle, 210

Parabola, 171
Paraboloid, 335
Parameter, auxiliary variable, 18
——, arbitrary constant, 155
Partial fractions, 133
—— derivatives, 334
—— differential equations, 490
Particle, (219)
Particular integral, 71, 455, 505
—— solution, 429
Pendulum, 8, (266), (417), 524
Piston pressure, (20), (240)
—— acceleration, (20), (311)
Perfect gas, 379, (487)
Period, 303
Permutations, arrangements, 89
Phase angle, 303
Pivot friction, 239
Plane curves, properties of, 357
——, standard equations, 331
—— tangent to surface, 346
—— normal to curve, 348
Polar coordinates, 149
—— subtangent and subnormal, 195
Pole, coordinates, 149
Polyphase relations, 307
Potential function, 382
Power, 236
—— series, 286
—— curves, logarithmic plotting, 420
Precision measure, 392
—— in computation, 411
Pressure, water, (87), 237
——, ——, center of, 238
Primitive of differential equation, 429
—— of partial differential equation, 493
Prismoid rule, 255
Probability, 385
—— integral, 392
Probable error, 396
Product of inertia, (248)
Projectile, (53), (210), (476)
Propagation of error, 413

Quadratic equations, 93
Quadric surfaces, 335

Radical axis of circles, 186
Radius of curvature, 214
—— of gyration, 228, 369
—— vector, 149
Range, of variable, 2
Rectangular hyperbola, 177
Redundant set of equations, 102
Rejection of observations, 411
Remainder, Taylor's theorem, 268
—— theorem, 88
Removable discontinuity, 34
Rolle's theorem, 200
Roots, rational, 91
——, exact, 93
——, approximate, 115
——, multiple, 201

Rotation of axes, 165
Rotor, 295

Salient point, 358
Semi-cubical parabola, (202), (216)
Semi-logarithmic plotting, 421
Separation of variables, 431
Series, Taylor's, 272
——, convergence, 274
——, alternating, 277
——, operations on 281
——, power, 286
——, Fourier, 311
——, solution of differential equations, 529
Shaft, (417), (454), (477), (512)
Shearing force, 237, 243
Short arc integration, 514
Simpson's rule, 251
Simultaneous equations, determinants, 101
—— ——, approximate methods, 121
—— differential equations, 469
—— total differential equations, 485
Singular points, 357
—— solution, differential equation, 445
—— ——, partial differential equation, 497
Slope of line, 157
Solution of equations, 88
—— of differential equation, 428
—— of partial differential equation, 493
Space coordinates, 328
Speed, 39, (52), (57), (65), (73), (82), (440)
Spirals, (215), (220)
Straight line, 157
—— —— in space, 329
Stress, 237, 238
Subnormal, 193
Subtangent, 192
Successive derivatives, 66
—— integration, 241
Surface, area of, 371
——, helicoidal, 336
——, quadric, 335
—— of revolution, area, 223
—— —— ——, volume, 225
—— —— ——, equation, 334
—— —— ——, parametric form, 336
Symbolic representation, harmonics, 306
—— operator, 459
Symmetry of curve, 152
—— of surfaces, 334
Synthetic division, 92

Tacnode, 358
Tangent line, conic sections, 187
—— ——, plane curves, 192
—— ——, space curves, 348
—— plane, surfaces, 346
Taylor's theorem, 267
—— ——, extended, 348
—— series, 272
—— ——, several variables, 350
Tensor, 295
Total derivative, 339
—— differential, 340
—— —— equations, 479
Traces of a surface, 334
Tractrix, (19), (441)
Transcendental functions, definition, 4
—— ——, graphs, 8
Transformation of coordinates, 163
Translation of axes, 165
Trapezoidal rule, 249
Trial integral, 131, 455
Trigonometric functions, graphs, 8
—— ——, derivatives, 53
—— ——, complex variable, 298
—— equations, 107
Triple integrals, 366

Unbounded variation, 28
Unilateral limits, 24

Variation of parameters, 475
Vectorial angle, 149
Vertex, of parabola, 172
Volumes, simple integrals, 225
——, approximate, 255
——, tetrahedron, 332
——, multiple integrals, 367

Weighted observations, 398
—— mean, 399
—— observation equations, 402
Work, (81), (82), 236, 381, (488)

Sans Tache

Sans Tache

IN THE "elder days of art" each artist or craftsman enjoyed the privilege of independent creation. He carried through a process of manufacture from beginning to end. The scribe of the days before the printing press was such a craftsman. So was the printer in the days before the machine process. He stood or fell, as a craftsman, by the merit or demerit of his finished product.

Modern machine production has added much to the worker's productivity and to his material welfare; but it has deprived him of the old creative distinctiveness. His work is merged in the work of the team, and lost sight of as something representing him and his personality.

Many hands and minds contribute to the manufacture of a book, in this day of specialization. There are seven distinct major processes in the making of a book: The type must first be set; by the monotype method, there are two processes, the "keyboarding" of the MS and the casting of the type from the perforated paper rolls thus produced. Formulas and other intricate work must be hand-set; then the whole brought together ("composed") in its true order, made into pages and forms. The results must be checked by proof reading at each stage. Then comes the "make-ready" and press-run and finally the binding into volumes.

All of these processes, except that of binding into cloth or leather covers, are carried on under our roof.

The motto of the Waverly Press is *Sans Tache.* Our ideal is to manufacture books *"without blemish"*—worthy books, worthily printed, with worthy typography—books to which we shall be proud to attach our imprint, made by craftsmen who are willing to accept open responsibility for their work, and who are entitled to credit for creditable performance.

The printing craftsman of today is quite as much a craftsman as his predecessor. There is quite as much discrimination between poor work

and good. We are of the opinion that the individuality of the worker should not be wholly lost. The members of our staff who have contributed their skill of hand and brain to this volume are:

Keyboard: Katharine Kocent, Harry Susemihl, Mildred Lambert.

Casters: Kenneth Brown, Charles Aher, Ernest Wann, Mahlon Robinson, George Smith, Frank Malanosky, Martin Griffen, Henry Lee, Charles Fick, George Bullinger.

Proof Room: Sarah Katzin, Mary Reed, Alice Reuter, Ruth Treischman, Ethel Strasinger, Dorothy Strasinger, Lucile Bull, Angeline Eifert, George Southworth, Ida Zimmerman, Audrey Tanner, Lillian Gilland.

Composing Room: Henry Shea, Nathan Miller, Harry Harmeyer, Andrew Rassa, Steve Simmons, John Flanagan, Anthony Wagner, Harry LaMotte, James Jackson, William Sanders, Austin Uhland, Ernest Salgado, Ray Kauffman, Arthur Baker, William Fite, George Moss, Edward Rice, Richard King, Theodore Nilson.

Press Room: Hugh Gardner, August Hildebrand.

Folder: Laurence Krug, Shipley Dellinger.

Cutter: William Armiger.

For the Mathematician

INTERPOLATION

By

J. F. STEFFENSEN

University of Copenhagen

Views the whole subject of finite differences, interpolation and mechanical quadration from a more general point of view than books of the past and places it in its proper niche of our present day mathematical literature.

Presents material equally applicable to the needs of the actuary, the statistician, the engineer, the astronomer, the physicist and the biologist, as well as to the student of pure mathematics.

A thorough treatment, yet it requires no more than an elementary knowledge of differential and integral calculus for its understanding.

Cloth, gold stamped *6 x 9* *248 pages*
47 tables, 5 figures

$8.00

THE WILLIAMS & WILKINS COMPANY

Publishers of Scientific Books and Periodicals

BALTIMORE, U. S. A.

BENEDICTINE UNIVERSITY LIBRARY
510R67M C001
THE MATHEMATICS OF ENGINEERING BALTIMOR
3 2811 00041 3601